Convergence to equilibrium for finite Markov processes, with application to the Random Energy Model.

Pierre Mathieu [1] and Pierre Picco [2].

CMI and CPT–CNRS, Marseille

Abstract: *We estimate the distance in total variation between the law of a finite state Markov process at time t, starting from a given initial measure, and its unique invariant measure. We derive upper bounds for the time to reach the equilibrium. As an example of application we consider a special case of finite state Markov process in random environment: the Metropolis dynamics of the Random Energy Model. We also study the process of the environment as seen from the process.*

Key Words: *Finite Markov processes, Spectral gap, Poincaré inequality, Random Spin Systems, Spin glasses, Metropolis Dynamics.*

AMS Classification Numbers: 60K35, 82B44, 82D30, 82C44

Abbreviated title: Finite Markov processes.

[1] CMI, Université de Provence, 39 Rue F. Joliot Curie, 13453 Marseille Cedex 13, France. pmathieu@gyptis.univ-mrs.fr

[2] CPT. CNRS Luminy, case 907, 13288 Marseille Cedex 9, France. Picco@cpt.univ-mrs.fr

I. Introduction

We are interested in estimating the speed of convergence towards equilibrium for a finite and reversible Markov chain, a well studied problem in the theory of Markov chains, see [11] for instance. Most, if not all results in this direction yield bounds on the distance to equilibrium which are uniform with respect to the initial distribution of the chain. In this paper, we shall rather derive estimates on mixing times that take into account the dependence on the initial law. As an example of application of our method, we study the Metropolis dynamics of Derrida's Random Energy Model (REM.).

Convergence times for the Metropolis dynamics of spin glasses were considered in [7]. Let us note that the present paper was done simultaneously with [7] and quoted therein as [11] with a slightly different title. In [7], estimates on the convergence time that depend on the initial law are given for models of spin glasses such as the REM or the Sherrington-Kirkpatrick model at high temperature. Three dynamics are considered: the random hoping time dynamics (RHT), the Glauber dynamics and the Metropolis dynamics. The initial configuration of the dynamics is always assumed to be chosen uniformly among all configurations.

To compare the results obtained in the two articles, let us mention that the starting points of the present article and [7] are the same: the generalized Poincaré inequalities that were introduced in [5], see section II here and in [7]. However the way to estimate the associated constant $\mathcal{L}_\eta(p)$, see (2.7) here and (2.2) there, are completely different. We will come back to this point later.

Since two slightly different notions of convergence time are used here and there, we first note that in [7], the time called $T^\omega(c)$, is defined as in (2.8), with c playing the rôle of ϵ. $T^\omega(c)$ a priori depends on the realizations of the energies as the ω emphasizes. In any case, the initial law, η, is uniform.

Here, for the Metropolis dynamics of the REM, the results are given in term of a time denoted $T_N(\epsilon, c, \eta)$ which is independent of the realizations of ω, see (4.11). It follows from the definition (4.11) that on a subset Ω_N of realizations of energies that has a probability larger than $1 - e^{-cN}$ we have

$$T_N(\epsilon, c, \eta) \geq T^\omega(\epsilon)$$

in particular this implies that, almost surely

$$\limsup_{N\to\infty} \frac{1}{N} \log T^\omega(\epsilon) \leq \limsup_{N\to\infty} \frac{1}{N} \log T_N(\epsilon, c, \eta) \tag{1.1}$$

We now recall some results from [4] and [7] for the convergence time of the Metropolis dynamics of the REM. In [7], it was proven that for η the uniform measure on $\{-1,+1\}^N$,

we have, for almost all ω

$$\limsup_{N\to\infty} \frac{1}{N} \log T^{\omega}(\epsilon) \leq 2\beta^2 \text{ when } \beta \leq \beta_c \tag{1.2}$$

and

$$\limsup_{N\to\infty} \frac{1}{N} \log T^{\omega}(\epsilon) \leq 2\beta\beta_c \text{ when } \beta \geq \beta_c \tag{1.3}$$

(Remember that the free energy and the mean energy per site converge for almost all ω as it follows from [8]). Note however that using the spectral gap estimates for the Metropolis Dynamics of the REM given in [4], we immediately get that, for all $\beta > 0$, almost surely in ω

$$\limsup_{N\to\infty} \frac{1}{N} \log T^{\omega}(\epsilon) \leq \beta\beta_c \tag{1.4}$$

and by checking all the probability estimates in [4], we also have for all $\beta > 0$, for all $c > 0$

$$\limsup_{N\to\infty} \frac{1}{N} \log T_N(\epsilon, c, \eta) \leq \beta\beta_c \tag{1.5}$$

Therefore, (1.2) gives an better estimate than (1.4) only for $\beta \leq \beta_c/2$. For $\beta > \beta_c/2$, (1.4) gives an better estimate than (1.2) and (1.3).
Here we prove that for the Metropolis Dynamics of the REM, for all $\beta \leq \beta_c$,

$$\limsup_{N\to\infty} \frac{1}{N} \log T_N(\epsilon, c, \eta) \leq \beta^2 \tag{1.6}$$

which together with (1.1) and (1.5) gives for all $\beta > 0$ a better estimate than (1.2) and (1.3). Thus we have improved the results of [7] in two ways: first we are using a more precise definition for the convergence time, second we won a factor 2 in the upper bound for $\beta \leq \beta_c$.
Note however that to get (1.4) or (1.5) a very careful analysis of optimization problems for paths on the weighted graph structure induced by the transition matrix of the dynamics was used. To prove (1.6), a similar analysis is needed. Thus, using the specific paths constructed in [4], instead of techniques based on estimates of the partition function as in [7], leads, for the Metropolis dynamics of the REM, to an improvement by a factor 2 in the estimates.
We believe that the bound (1.6) is sharp $i.e$ $\lim_{n\uparrow\infty} \frac{1}{N} \log T_N(\epsilon, c, \eta) = \beta^2$ for $\beta \leq \beta_c$.
We also believe that a similar analysis could be carried over for the Glauber dynamics, but the numerical factor in front of β^2 in (1.6) would then be different. As far as the Random Hoping Time dynamics is concerned, it seems that the techniques of [7] directly

lead to upper bounds of the correct order. Note however that the RHT dynamics has a much simpler structure than the Metropolis or Glauber ones. Indeed the RHT dynamics is nothing but a time-changed standart random walk on configuration space. The sequence of the different states visited by the process is independent of the Hamiltonian. On one hand, this feature very much simplifies the geometry. On the other hand, physicists believe that the evolution of the process should rather look like a random perturbation of the steepest gradient dynamical system. The RHT dynamics displays un-physical features.

The organization of the paper is as follows: part II and II deal with general reversible Markov chains on a finite set. In part II, we define generalized Poincaré inequalities and show that they control the decay of the semi-group (Theorem 2.1). Then we derive geometric estimates for the generalized Poincaré constants (Theorem 2.2). Part III contains an application of these results in a case where the state space can be splitted into two components: 'good' and 'bad' points. For the reader's convenience, we decided to give self-contained proofs of our results at the risk of repeating arguments already used in [5], [6] or [7] .

Although we shall not directly use the results of part III to study the R.E.M., the strategy will be the same. Only technical aspects make the computation for the R.E.M. a little longer than the proof in part III. In part IV, we precisely define the R.E.M. and state our bounds for the thermalization time (Theorems 4.1 and 4.2). Then we proceed to the proofs. In part V, we extend our results to the process of the environment as seen from the particle. This section is similar to the section 3 of [7] with more pedagogical details on the construction of the process. We then show that the equilibrium time also satisfies (4.6). Part VI contains the proof of some static estimates on the R.E.M. that we needed in the previous parts.

II. Generalized Poincaré inequalities

Let $X = (X_t)_{t\geq 0}$ be an homogeneous Markov process on a finite state space, $\mathcal{X}$. We assume that there is a unique invariant, ergodic probability measure for X, say π. We further assume that π charges every point in $\mathcal{X}$ and that it is reversible. Let η be some probability measure on $\mathcal{X}$ and call $\mathcal{L}_\eta(X_t)$ the law of X_t when the initial law is η. We wish to bound $d_{TV}(\mathcal{L}_\eta(X_t), \pi)$, the distance in total variation between the law of X at time t and the equilibrium law π. More precisely, we would like to obtain an upper bound in terms of the geometry of the Markov process X i.e. in terms of the geometry of the graph structure induced by the transition matrix on the state space.

It is well known that one can use Poincaré inequalities to bound $d_{VT}(\mathcal{L}_\eta(X_t), \pi)$. Indeed calling λ the spectral gap of the generator of X (which is a symmetric matrix since we

have assumed that π is reversible), we have, for any real valued function f defined on $\mathcal{X}$ and for any $t \geq 0$,

$$\pi[(P_t f - \pi(f))^2] \leq e^{-2\lambda t} \pi(f^2) \tag{2.1}$$

where P_t denotes the semi-group i.e. $P_t f(x) = E_x[f(X_t)]$. From (2.1), it immediately follows that

$$max_{x \in \mathcal{X}} d_{TV}(\mathcal{L}_x(X_t), \pi) \leq \sqrt{\frac{1}{\pi_*}} e^{-\lambda t} \tag{2.2}$$

where $\mathcal{L}_x(X_t)$ is the law of X_t when the initial law is a Dirac mass at the point $x \in \mathcal{X}$ and $\pi_* = \min_{x \in \mathcal{X}} \pi(x)$. It now remains to estimate λ in terms of the geometry of X. Such bounds exist, they rely on Poincaré inequalities: assume that for some constant $a > 0$ and any function f with $\pi(f) = 0$, we have:

$$\pi(f^2) \leq a \mathcal{E}(f, f) \tag{2.3}$$

then $1/\lambda \leq a$. Here $\mathcal{E}$ is the Dirichlet form of X. From (2.3) one can deduce lower bounds of λ in terms of optimization problems for paths on the weighted graph structure induced by the transition matrix of X on $\mathcal{X}$ (See [11] and the references therein). (2.2) might be sharp or not depending on X. Many efforts were recently made to improve (2.2). More precise bounds can be obtained replacing the Poincaré inequality by more sophisticated functional inequalities such as Log-Sobolev, Sobolev or Nash inequalities. We refer to [11] for a detailed discussion of this topic. In all cases, one estimates $max_{x \in \mathcal{X}} d_{TV}(\mathcal{L}_x(X_t), \pi)$ i.e. the speed of convergence to equilibrium starting from the worst initial point.

We look for estimates of $d_{TV}(\mathcal{L}_\eta(X_t), \pi)$ that should depend on η. This paper is an attempt to adapt the strategy of the Poincaré inequality in this context: for each initial law η, we introduce a family of functional inequalities, quite similar to the Poincaré one, and prove that they allow one to control the distance to equilibrium. We call these inequalities *generalized Poincaré inequalities*. We then derive geometric bounds for the constants involved in these inequalities in the spirit of [11].

Let $(K(x, y), (x, y) \in \mathcal{X}\text{x}\mathcal{X})$ be the transition matrix of the Markov process X. Since we assume that the measure π is reversible, the kernel $k(x, y) = K(x, y)/\pi(x)$ is symmetric, i.e. $k(x, y) = k(y, x)$. Let $P_t f(x) = E_x[f(X_t)]$ denote the semi-group associated to X.

For functions f and g defined on $\mathcal{X}$, let

$$\mathcal{E}(f, g) == \frac{1}{2} \sum_{x,y} (f(x) - f(y))(g(x) - g(y)) k(x, y) \pi(x) \pi(y) \tag{2.4}$$

be the Dirichlet form of X. For any edge $e = (x, y) \in \mathcal{X}\text{x}\mathcal{X}$, let $Q(e) = k(x, y)\pi(x)\pi(y)$. Also define $d_e f = f(x) - f(y)$. Then (2.4) can be re-written as

$$\mathcal{E}(f, g) = \frac{1}{2} \sum_{e \in \mathcal{X}^2} Q(e) d_e f d_e g \tag{2.5}$$

For $p \in]0,1]$, let us define the following constants:

$$\mathcal{L}(p) = \inf_{f \ s.t. \ \pi(f)=0} \frac{\mathcal{E}(f,f)\|f\|_\infty^{(2-2p)/p}}{\pi(|f|)^{2/p}} \tag{2.6}$$

and, for a probability measure on $\mathcal{X}$, say η,

$$\mathcal{L}_\eta(p) = \inf_{f \ s.t. \ \pi(f)=0} \frac{\mathcal{E}(f,f)\|f\|_\infty^{(2-2p)/p}}{\eta(|f|)^{2/p}} \tag{2.7}$$

Clearly $\mathcal{L}(p) = \mathcal{L}_\pi(p)$. Hölder's inequality implies that the function $p \to \mathcal{L}_\eta(p)$ is decreasing and that $\mathcal{L}(p) \geq \lambda$ for any p. (Remember that λ denotes the spectral gap of the generator of X).

To measure the time it takes for the process to reach equilibrium, we define the following quantities:

$$d_\eta(t) = \sup_{s \geq t} \sup_{f; \|f\|_\infty \leq 1} \eta(|P_s f - \pi(f)|)$$

and, for any $\varepsilon > 0$,

$$T_\eta(\varepsilon) = \inf\{t > 0 \ s.t. \ d_\eta(t) \leq \varepsilon\} \tag{2.8}$$

Note that $d_{TV}(\mathcal{L}_\eta(X_s), \pi) \leq d_\eta(t)$ for all $s \geq t$.

Remark: let

$$\Lambda(p) = \inf_{f \ s.t. \ \pi(f)=0} \frac{\mathcal{E}(f,f)}{\pi(|f|^p)^{2/p}}$$

Then, as a consequence of Hölder's inequality, $\Lambda(p) \leq \mathcal{L}(p)$ for $p \in]0,1]$. Also $\Lambda(1) = \mathcal{L}(1)$. The constants $\Lambda(p)$ and $\mathcal{L}(p)$ have already been introduced in [5]. (In the notation of [5], $\mathcal{L}(p)$ is denoted $\mathcal{K}(p/(1-p), +\infty)$). It follows from the results of [5], that $\mathcal{L}(p)$ can also be defined in terms of the capacity associated to $\mathcal{E}$ and different estimates of hitting times can be derived in terms of $\mathcal{L}(p)$.* We also have $\Lambda(2) = \lambda$ and $\Lambda(p) \geq \lambda$ for any $p \in]0,1]$. Because of the similarity of the definition of $\Lambda(p)$ and the Poincaré inequality, we call the inequality $\Lambda(p) \geq a$ for some $a > 0$, a "generalized Poincaré inequality", although there is no spectral interpretation.

Theorem 2.1 : *let $p \in]0,1]$ and $p' \in]0,1]$. There exists a universal function of (p,p'), $C_{p,p'}$, such that, for any probability measure η and any $t > 0$,*

$$d_\eta(t) \leq C_{p,p'} \mathcal{L}_\eta(p')^{-p'/2} \mathcal{L}(p)^{-pp'/(4-2p)} t^{-p'/(2-p)} \tag{2.9}$$

$C_{p,p'} = e^{-p'/2}(p/(2-p))^{pp'/(4-2p)}$ would do. As a consequence, for any $\varepsilon > 0$, we have

$$T_\eta(\varepsilon) \leq \tilde{C}_p \mathcal{L}_\eta(p')^{-(2-p)/2} \mathcal{L}(p)^{-p/2} \varepsilon^{-(2-p)/p'} \tag{2.10}$$

* We take this opportunity to warm the reader that the results of part II in [5] are false.

where $\tilde{C}_p = e^{-(2-p)/2}(p/(2-p))^{p/2}$.

Proof: : we shall prove that, for any function f with $\pi(f) = 0$, then

$$\eta(|P_t f|) \leq C_{p,p'} \mathcal{L}_\eta(p')^{-p'/2} \mathcal{L}(p)^{-pp'/(4-2p)} t^{-p'/(2-p)} \|f\|_\infty \tag{2.11}$$

with $C_{p,p'} = e^{-p'/2}(p/(2-p))^{pp'/(4-2p)}$. (2.11) implies (2.9).

Step 1: define

$$\mathcal{K}(p) = \inf_{f \ s.t. \ \pi(f)=0} \frac{\mathcal{E}(f,f)\|f\|_\infty^{(4-2p)/p}}{\pi(|f|^2)^{2/p}} \tag{2.12}$$

From Hölder's inequality, we deduce that $\mathcal{K}(p) \geq \mathcal{L}(p)$. Let f be s.t. $\pi(f) = 0$. We claim that

$$\pi[(P_t f)^2] \leq (\frac{4-2p}{p})^{-p/(2-p)} (\mathcal{K}(p)t)^{-p/(2-p)} \|f\|_\infty^2 \tag{2.13}$$

Then (2.13) will also hold with $\mathcal{K}(p)$ replaced by $\mathcal{L}(p)$.

Proof of (2.13): let $f(t) = \pi[(P_t f)^2]$. Then $f'(t) = -2\mathcal{E}(P_t f, P_t f)$. By definition of $\mathcal{K}(p)$, we have:

$$f'(t) \leq -2\mathcal{K}(p) \frac{f(t)^{2/p}}{\|P_t f\|_\infty^{(4-2p)/p}}$$

Since P_t is a contraction in L_∞, we also have

$$f'(t) \leq -2\mathcal{K}(p) \frac{f(t)^{2/p}}{\|f\|_\infty^{(4-2p)/p}}$$

Integrating this last inequality, we get

$$\begin{aligned} f(t)^{1-2/p} &\geq f(0)^{1-2/p} + \frac{4-2p}{p} t \frac{\mathcal{K}(p)}{\|f\|_\infty^{(4-2p)/p}} \\ &\geq \frac{4-2p}{p} t \frac{\mathcal{K}(p)}{\|f\|_\infty^{(4-2p)/p}} \end{aligned}$$

which implies (2.13). ■

Step 2: there exists a universal constant C s.t. for any function f and any $t > 0$ we have

$$\mathcal{E}(P_t f, P_t f) \leq (C/t)\pi[f^2] \tag{2.14}$$

($C = 1/(2e)$ would do.)

Proof: for all $\mu \geq 0$ and $t > 0$, we have $\mu e^{-2\mu t} \leq C/t$. Use this inequality and a spectral decomposition of the Dirichlet form $\mathcal{E}$ to deduce (2.14). ■

Step 3: we finish the proof of (2.11). By definition of $\mathcal{L}_\eta(p')$, we have:

$$\eta(|P_t f|)^{2/p'} \le \mathcal{L}_\eta(p')^{-1}\mathcal{E}(P_t f, P_t f)\|P_t f\|_\infty^{(2-2p')/p'}$$

Using (2.14), the semi-group property: $P_t = P_{t/2}P_{t/2}$, and the fact that P_t is a contraction in L_∞, we get that

$$\eta(|P_t f|)^{2/p'} \le C\frac{2}{\mathcal{L}_\eta(p')t}\pi(|P_{t/2}f|^2)\|f\|_\infty^{(2-2p')/p'}$$

Using (2.12) (with $\mathcal{L}(p)$ instead of $\mathcal{K}(p)$), we get that

$$\eta(|P_t f|)^{2/p'} \le 2C(\frac{4-2p}{p})^{-p/(2-p)}(\mathcal{L}_\eta(p')t)^{-1}(\frac{2}{\mathcal{L}(p)t})^{p/(2-p)}\|f\|_\infty^{2/p'}$$

■

Remarks:

(i) Depending on the concrete example under consideration, the sharpness of the bound (2.9) ranges from good to extremely bad. Let us just outline one example where Theorem 2.1 leads to a very bad estimate: we consider the usual random walk on the discrete cube $\mathcal{X} = \{-1,+1\}^N$. Then $Q(e) = 1/(N2^N)$, for any edge between two nearest neighbours in $\mathcal{X}$. Choose for η a Dirac mass, say $\eta = \delta_a$. Using the test function $f = \delta_a - \pi(a)$ in formula (2.7), we get that, for large enough N,

$$\mathcal{L}_\eta(p) \le 2^{2/p-N}$$

Therefore (2.10) would lead to the conclusion that the process reaches equilibrium in a time shorter than $\exp(cN)$, whereas the true value of $T_\eta(\varepsilon)$ is known to be of order $N\log N$. We will see with the R.E.M. an example where Theorem 2.1 leads to more interesting conclusions.

There is one situation in which (2.9) is not so far from being sharp: assume that $\eta = \pi$. Let a be such that, for any function f with $\pi(f) = 0$, and for any time $t > 0$, we have

$$\pi(|P_t f|] \le (\frac{a}{t})^{\frac{p}{2-p}}\|f\|_\infty \tag{2.15}$$

By interpolation, (2.15) implies that

$$\pi[(P_t f)^2] \le (\frac{a}{t})^{\frac{p}{2-p}}\|f\|_\infty^2 \tag{2.16}$$

Use now the inequality

$$\pi[f^2] - \pi[(P_t f)^2] = \int_0^t 2\mathcal{E}(P_s f, P_s f)ds \le 2t\mathcal{E}(P_t f, P_t f)$$

to get that

$$\pi(f^2) \leq 2t\mathcal{E}(P_t f, P_t f) + (\frac{a}{t})^{\frac{p}{2-p}} \|f\|_\infty^2$$

Choosing the best value for t, we obtain the inequality:

$$\pi(f^2) \leq C_p a^{\frac{p}{2}} \|f\|_\infty^{2-p} (\mathcal{E}(f,f))^{\frac{p}{2}}$$

,where C_p is some universal function of p. In other words we have proved that $1/\mathcal{K}(p) \leq C_p a$, i.e (2.13) is sharp, up to multiplicative constants.
(ii) We derive estimates of the eigenvectors of $\mathcal{E}$ in terms of $\mathcal{L}(p)$. Following the terminology of [5], let us define

$$\mathcal{K}_2(p) = \inf_{f \ s.t. \ \pi(f)=0} \frac{\mathcal{E}(f,f)\pi(f^2)^{(1-p)/2p}}{\pi(|f|)^{(p+1)/p}} \tag{2.17}$$

It follows from Proposition 1, Proposition 2 and Theorem 1 in [5] that, for any $p' < p$, there exists a constant $C_{p,p'}$ such that $\mathcal{L}(p) \leq C_{p,p'}\mathcal{K}_2(p')$.
Let now l be an eigenvalue of $\mathcal{E}$ and ϕ be the corresponding eigenvector. We assume that $l \neq 0$ (ϕ is not constant), and $\pi[\phi^2] = 1$. Using $f = \phi$ in (2.17) and $\mathcal{E}(\phi,\phi) = l$, we obtain that $\mathcal{K}_2(p) \leq l/\pi(|\phi|)^{(p+1)/p}$. Replacing $\mathcal{K}_2(p)$ by $\mathcal{L}(p)$, we therefore have:

$$\pi(|\phi|) \leq C_{p,p'}(\frac{l}{\mathcal{L}(p)})^{p'/(1+p')} \tag{2.18}$$

for any $p' < p$.
(2.18) implies that, if l is much smaller than $\mathcal{L}(p)$, then $\pi(|\phi|)$ is small i.e. the function ϕ is very concentrated on its support. Since $l \geq \lambda$, where λ is the spectral gap, this situation can occur only if, for some p, $\lambda << \mathcal{L}(p)$. This will be the case for Metropolis dynamics of the R.E.M. at high temperature and we shall use (2.18) to prove that the first eigenvector of the dynamics is degenerate.
Geometric estimates: a path γ in $\mathcal{X}$ is a sequence of vertices $\gamma = (x_0, ..., x_k)$. Equivalently, γ can be viewed as a sequence of bounds $\gamma = (e_1, ..., e_k)$ with $e_i = (x_{i-1}, x_i)$. The length of γ is $|\gamma| = k$. For $x, y \in \mathcal{X}$, let $\Gamma(x,y)$ be the set of all paths $\gamma = (x_0, ..., x_k)$ with $x_0 = x$ and $x_k = y$ and $k(x_{i-1}, x_i) \neq 0$ for all $i = 1...k$. For each $x \neq y \in \mathcal{X}$, let us choose one path, say $\gamma(x,y) \in \Gamma(x,y)$. Since we have assumed that π is ergodic and charges all points in $\mathcal{X}$, X is irreducible and therefore $\Gamma(x,y)$ is always non empty.

Theorem 2.2 : *(i) Let $p \in]0,1[$. Let $\lambda(x)$ and $\mu(x)$ be two positive functions on $\mathcal{X}$. We have*

$$\frac{1}{\mathcal{L}_\eta(p)} \leq 2^{2/p-1} \left(\sum \pi(x)\lambda(x)^{p/(1-p)} \sum \eta(y)\mu(y)^{p/(1-p)} \right)^{(2-2p)/p}$$
$$\left(\sum_{e \ s.t. \ Q(e)\neq 0} \frac{1}{Q(e)} (\sum_{x,y \ s.t. \ e\in\gamma(x,y)} \frac{\pi(x)\eta(y)}{\lambda(x)\mu(y)})^2 \right) \tag{2.19}$$

(ii)

$$\frac{1}{\mathcal{L}_\eta(1)} \leq 2\left(\sum_{e\ s.t.\ Q(e)\neq 0} \frac{1}{Q(e)}(\sum_{x,y\ s.t.\ e\in\gamma(x,y)} \pi(x)\eta(y))^2\right) \tag{2.20}$$

Comments: let us recall from [11] the following estimate of the spectral gap:

$$\frac{1}{\lambda} \leq \max_{e\ s.t.\ Q(e)\neq 0} \frac{1}{Q(e)} \sum_{x,y\ s.t.\ e\in\gamma(x,y)} |\gamma(x,y)|\pi(x)\pi(y) \tag{2.21}$$

Proof: : (ii) follows from (i): choose $\lambda(x) = \mu(x) = 1$, and let p tend to 1.
Let f be a function s.t. $\pi(f) = 0$. Note that $f(y) - f(x) = \sum_{e\in\gamma(x,y)} d_e f$. Therefore

$$\begin{aligned}
\Sigma_y\eta(y)|f(y)| =& \Sigma_y\eta(y)|f(y)| - \Sigma_x\pi(x)f(x)\\
&\leq \Sigma_{x,y}\eta(y)\pi(x)|f(x) - f(y)|\\
&= \Sigma_{x,y}\eta(y)\pi(x)|\Sigma_{e\in\gamma(x,y)}d_e f|\\
&\leq 2^{1-p}\Sigma_{x,y}\eta(y)\pi(x)|\Sigma_{e\in\gamma(x,y)}d_e f|^p\|f\|_\infty^{1-p}\\
&= 2^{1-p}\|f\|_\infty^{1-p}\Sigma_{x,y}\eta(y)\pi(x)[\lambda(x)\mu(y)|\Sigma_{e\in\gamma(x,y)}d_e f|]^p\lambda(x)^{-p}\lambda(y)^{-p}
\end{aligned}$$

We apply Hölder's inequality to get

$$\begin{aligned}
&\Sigma_y\eta(y)|f(y)|\\
&\leq 2^{1-p}\|f\|_\infty^{1-p}\left(\Sigma_{x,y}\pi(x)\eta(y)\lambda(x)^{p/(1-p)}\mu(y)^{p/(1-p)}\right)^{1-p}\left(\Sigma_{x,y}\frac{\pi(x)\eta(y)}{\lambda(x)\mu(y)}|\Sigma_e d_e f|\right)^p\\
&\leq 2^{1-p}\|f\|_\infty^{1-p}\left(\Sigma_{x,y}\pi(x)\eta(y)\lambda(x)^{p/(1-p)}\mu(y)^{p/(1-p)}\right)^{1-p}\times\\
&\qquad\times\left(\Sigma_e|d_e f|\Sigma_{x,y\ s.t.\ e\in\gamma(x,y)}\frac{\pi(x)\eta(y)}{\lambda(x)\mu(y)}\right)^p
\end{aligned}$$

Applying once more Hölder's inequality, we get

$$\begin{aligned}
&\Sigma_y\eta(y)|f(y)|\\
&\leq 2^{1-p}\|f\|_\infty^{1-p}\left(\Sigma_{x,y}\pi(x)\eta(y)\lambda(x)^{p/(1-p)}\mu(y)^{p/(1-p)}\right)^{1-p}\times\\
&\qquad\times\left(\Sigma_e(d_e f)^2Q(e)\right)^{p/2}\left(\Sigma_e\frac{1}{Q(e)}(\Sigma_{x,y\ s.t.\ e\in\gamma(x,y)}\frac{\pi(x)\eta(y)}{\lambda(x)\mu(y)})^2\right)^{p/2}
\end{aligned} \tag{2.22}$$

Replacing $\Sigma_e Q(e)|d_e f|^2$ by $2\mathcal{E}(f,f)$, we get the desired result. ■

III. Applications

This part of the paper mainly has a pedagogical aim. We shall illustrate how one can use the results of part II in a concrete situation. An even more concrete example of application will be given in the next part with the R.E.M.

Comparing (2.20) and (2.21), one sees that the gain in using generalized spectral gap inequalities instead of the usual spectral gap inequality is that we can now afford having some "very bad sites" since we replaced a "max" over edges e by a sum. Besides formula (2.19) gives us the possibility of 'killing' these bad points by choosing λ and μ. To illustrate the way it works, let us assume that the state space $\mathcal{X}$ can be divided into two disjoint sets, B and G. 'B' stands for 'bad'. Points in B are supposed to be pathological and we do not expect them to play any role on the speed of convergence when the initial measure is smooth enough.

The next Theorem states a lower bound for $\mathcal{L}_\eta(p)$ which is valid for any partition of $\mathcal{X}$ into two sets B and G, but (3.1) is useful only if, firstly, we assume that the measure of B is small both for π and η and besides we also assume somehow that the hitting time of B is large i.e. the weights $Q(e)$ for those edges e that touch B are not too small.

Let us introduce some notation:

$$\gamma^* = \sup_{x,y\in\mathcal{X}} |\gamma(x,y)|$$

$$\mathcal{B} = \{e \in \mathcal{X}\mathrm{x}\mathcal{X} \textit{ s.t. there exist } x \textit{ and } y \textit{ s.t. } e \in \gamma(x,y) \textit{ and } x \in B \textit{ or } y \in B\}$$

In $\mathcal{B}$ are edges $e \in \gamma(x,y)$ with both x and y in B.

Theorem 3.1 : *for any $p \in]0,1]$, for any probability measure η*

$$\begin{aligned}\frac{1}{\mathcal{L}_\eta(p)} \leq 2^{6/p-3}\{\gamma^* \sup_{e \ s.t.\ Q(e)\neq 0}\left(\frac{1}{Q(e)}\sum_{x\in G, y\in G \ s.t.\ e\in\gamma(x,y)}\pi(x)\eta(y)\right) \\ +2\left(\sum_{e\in\mathcal{B}}\frac{1}{Q(e)}\right)\left(\pi(B)^{2/p}+\eta(B)^{2/p}\right)\}\end{aligned} \tag{3.1}$$

Proof: : let $p \in]0,1[$. The proof for $p=1$ is simpler and we leave it to the reader. Let us choose λ and μ as follows: $\lambda(x) = \mu(x) = 1$ for $x \in G$, $\lambda(x) = \pi(B)^{1-1/p}$ for $x \in B$ and $\mu(x) = \eta(B)^{1-1/p}$ for $x \in B$. Then

$$\begin{aligned}&\sum \pi(x)\lambda(x)^{p/(1-p)} \\ =&\pi(G)+1 \leq 2\end{aligned}$$

The same holds for $\sum \eta(y)\mu(y)^{1/(1-p)}$. Therefore

$$\frac{1}{\mathcal{L}_\eta(p)} \leq 2^{6/p-5} \sum_e \frac{1}{Q(e)} \Big(\sum_{x,y \ s.t. \ e \in \gamma(x,y)} \frac{\pi(x)\eta(y)}{\lambda(x)\mu(y)} \Big)^2 \tag{3.2}$$

We compute the sum in (3.2) considering separately the cases $(x, y) \in G\text{x}G$, $(x, y) \in B\text{x}B$, $(x, y) \in G\text{x}B$ and $(x, y) \in B\text{x}G$. Since $\lambda = \mu = 1$ on G, the first term is bounded by

$$\begin{aligned}
&\sum_e \frac{1}{Q(e)} \Big(\sum_{x,y \in G \ s.t. \ e \in \gamma(x,y)} \pi(x)\eta(y) \Big)^2 \\
&\leq \left(\sup_e \frac{1}{Q(e)} \sum_{x,y \in G \ s.t. \ e \in \gamma(x,y)} \pi(x)\eta(y) \right) \left(\sum_e \sum_{x,y \in \mathcal{X} \ s.t. \ e \in \gamma(x,y)} \pi(x)\eta(y) \right) \\
&= \left(\sup_e \frac{1}{Q(e)} \sum_{x,y \in G \ s.t. \ e \in \gamma(x,y)} \pi(x)\eta(y) \right) \left(\sum_{x,y} |\gamma(x,y)| \pi(x)\eta(y) \right) \\
&\leq \left(\sup_e \frac{1}{Q(e)} \sum_{x,y \in G \ s.t. \ e \in \gamma(x,y)} \pi(x)\eta(y) \right) \gamma^*
\end{aligned} \tag{3.3}$$

The term corresponding to the case $(x, y) \in B\text{x}B$ is bounded by

$$\begin{aligned}
&\sum_{e \in \mathcal{B}} \frac{1}{Q(e)} \Big(\sum_{x,y \in B} \pi(x)\eta(y)\pi(B)^{1/p-1}\eta(B)^{1/p-1} \Big)^2 \\
&\leq \left(\sum_{e \in \mathcal{B}} \frac{1}{Q(e)} \right) \pi(B)^{2/p}\eta(B)^{2/p} \\
&\leq \left(\sum_{e \in \mathcal{B}} \frac{1}{Q(e)} \right) (\pi(B)^{2/p} + \eta(B)^{2/p})
\end{aligned} \tag{3.4}$$

The term corresponding to the case $(x, y) \in G\text{x}B$ is bounded by

$$\begin{aligned}
&\sum_{e \in \mathcal{B}} \frac{1}{Q(e)} \pi(G)^2 \eta(B)^{2/p} \\
\leq &\sum_{e \in \mathcal{B}} \frac{1}{Q(e)} \eta(B)^{2/p}
\end{aligned} \tag{3.5}$$

Similarly the contribution of $(x, y) \in B\text{x}G$ is bounded by

$$\sum_{e \in \mathcal{B}} \frac{1}{Q(e)} \pi(B)^{2/p} \tag{3.6}$$

Inserting these bounds in (3.2) leads to the statement of Theorem 3.1. ■

In the preceding Theorem, we chose the same 'bad' set for both measures π and η. We now describe a slightly more sophisticated version of Theorem 3.1 obtained when choosing a different bad set for π and η. Let us therefore assume that $\mathcal{X}$ can be split into the disjoint union of two sets B_η and G_η. B_η might differ from B. We modify the definition of $\mathcal{B}$ accordingly:

$$\mathcal{B} = \{e \in \mathcal{X}\text{x}\mathcal{X} \textit{ s.t. there exist } x \textit{ and } y \textit{ s.t. } e \in \gamma(x,y) \textit{ and } x \in B \textit{ or } y \in B_\eta\}$$

The proof of the following claim is identical to the proof of Theorem 3.1:

Theorem 3.2 : *for any* $p \in]0,1]$, *for any probability measure* η, *any partitions* $\mathcal{X} = B \cup G = B_\eta \cup G_\eta$, *we have*

$$\begin{aligned}\frac{1}{\mathcal{L}_\eta(p)} \leq 2^{6/p-3}\{\gamma^* \sup_{e \text{ s.t. } Q(e)\neq 0} \left(\frac{1}{Q(e)} \sum_{x\in G, y\in G_\eta \text{ s.t. } e\in\gamma(x,y)} \pi(x)\eta(y)\right) \\ + 2\left(\sum_{e\in\mathcal{B}} \frac{1}{Q(e)}\right)\left(\pi(B)^{2/p} + \eta(B_\eta)^{2/p}\right)\}\end{aligned} \tag{3.7}$$

Proof: : choose $\lambda(x) = 1$ for $x \in G$, $\mu(x) = 1$ for $x \in G_\eta$ and $\lambda(x) = \pi(B)^{1-1/p}$ for $x \in B$, $\mu(x) = \eta(B_\eta)^{1-1/p}$ for $x \in B_\eta$. Then proceed as in the proof of Theorem 3.1. ■

Finally let us mention that even more elaborated bounds can be obtained: we could distinguish bounds in $\mathcal{B}$ linking sites (x,y) with $(x,y) \in G\text{x}B_\eta$, $(x,y) \in B\text{x}G$ and $(x,y) \in B\text{x}B_\eta$. We could also introduce 'weights' on bounds. We could choose a 'flow' of paths rather that picking a single path from x to y. If necessary, one can also use these three tricks at the same time. We refer to Chapter 3 in [11] for the notions of 'weights' and 'flow' or even 'generalized weights'.

IV. Dynamical phase transition for the REM

Before stating our result, let us recall the definition and some known facts on the R.E.M.
Derrida's Random Energy Model: The REM was introduced by Derrida [1,2] as the simplest mean field spin glass. It is a caricature of the Sherrington & Kirkpatrick (SK) spin glass model [10]. Both are spin systems with Ising spins taking value ± 1. In the SK

model one has Gaussian pair interactions, while in the REM one has Gaussian multibody interactions of any order. The Hamiltonian of the REM is

$$H(\sigma) \equiv -\frac{\sqrt{N}}{2^{N/2}} \sum_{\alpha \subset \{1,\dots,N\}} J_\alpha \sigma_\alpha \tag{4.1}$$

where the sum is over all the 2^N subsets of $\{1,\dots,N\}$, $(J_\alpha, \alpha \subset \{1,\dots,N\})$ is a family of i.i.d. standard Gaussian variables defined on a common probability space $(\Omega, \Sigma, \mathcal{Q})$ and $\sigma_a \equiv \Pi_{i\in\alpha}\sigma_i$ with the convention that $\sigma_\emptyset = 1$. It turns out that the random variables $H(\sigma)$ and $H(\sigma')$ corresponding to different configurations $\sigma \neq \sigma'$ are independent Gaussian variables with zero mean and variance N. The equilibrium statistical mechanics of the REM has been well studied, e.g., in a non rigorous way, in [1,2] and, in a rigorous way, in [3,8,9]. We quote some of the (rigorous) results that will be important for understanding the dynamics. Given $\beta \geq 0$, the inverse temperature, let us denote by

$$Z_N \equiv Z_N(\beta) = \sum_\sigma e^{-\beta H(\sigma)} \tag{4.2}$$

the finite volume partition function and by

$$F_N(\beta) = \frac{1}{N} \log Z_N(\beta) \tag{4.3}$$

the finite volume free energy.
It was proved in [8] that for all $\beta \geq 0$ the limit $\lim_{N\to\infty} F_N(\beta) = F(\beta)$ exists $\mathcal{Q}$-almost surely and in $L^p(\Omega, \Sigma, \mathcal{Q})$ for $1 \leq p < \infty$. $F(\beta)$ equals $\beta^2/2 + \beta_c^2/2$ for $\beta < \beta_c$ and $\beta_c\beta$ for $\beta \geq \beta_c$, as expected from the results of [1]. $F(\beta)$ is therefore a non random function which is twice differentiable in β but the second derivative has a jump at $\beta_c = \sqrt{2\log 2}$. This is called in the physics literature a third order phase transition. Another important fact is that, depending wether we are in a high temperature regime ($\beta < \beta_c$) or in a low temperature one ($\beta \geq \beta_c$), not only does the free energy change from a quadratic function of β to a linear one but the difference between the finite volume free energy and its infinite volume limit is exponentially small in N in the high temperature case, whereas, in the low temperature regime, $F_N(\beta) - F(\beta)$ behaves as $C(\omega,\beta,N)\frac{\log N}{N}$, for some random function $C(\omega,\beta,N)$. $C(\omega,\beta,N)$ converges in $\mathcal{Q}$-probability to a non-random limit but does not converge $\mathcal{Q}$-almost surely and the $\mathcal{Q}$ almost-sure cluster set of $C(\omega,\beta,N)$ was identified in [9].
Let us now discuss the dynamical properties of the model. We consider the Metropolis dynamics. (See (4.8)). A first step in the study of the dynamics for the REM was done in [4]. There the spectral gap, λ_N of the usual single spin flip metropolis dynamics in volume

N is studied. In particular it was proved that for *all* inverse temperatures $\beta > 0$ we have, $\mathcal{Q}$-almost surely

$$\lim_{N\uparrow\infty} -\frac{1}{N}\log\lambda_N = \beta\beta_c \tag{4.4}$$

Moreover $\mathcal{Q}$-almost sure finite size corrections are also given in [4]: we have

$$\beta\beta_c - c\beta\sqrt{\frac{\log N}{N}} \le -\frac{1}{N}\log\lambda_N \le \beta\beta_c + c\beta\sqrt{\frac{\log N}{N}} \tag{4.5}$$

$\mathcal{Q}$-almost-surely, for all but a finite number of indices N, for some constant c.
However one would have expected the dynamics to present a kind of transition as the previously mentioned *static* phase transition that can be seen on the free energy $F(\beta)$. Such a *dynamical* transition is not seen on the spectral gap.
Thus we are lead to the following question: how can we see a dynamical phase transition on the single spin flip dynamics ?
The inverse spectral gap can be used as an estimate for the thermalization time of the dynamics. For the Metropolis dynamics, $1/\lambda_N$ is actually a sharp upper bound for the time it takes for the dynamics to reach equilibrium, whatever was the initial law. In particular we may consider the dynamics issued from a given configuration. The REM is rather pathological in the sense that the configurations of lowest energy (of order $-\beta_c N$) are surrounded (in a sense of a single spin flip) by configurations of energy of order at most $\pm\sqrt{N\log N}$. The bounds in (4.5) follow from this fact. Starting the dynamics at a configuration of lowest energy, we have to wait for a time of order $e^{N\beta\beta_c}$ before the first spin flip. As we see, the time to reach equilibrium starting from a configuration of minimal energy is therefore of order $e^{N\beta\beta_c}$.
In the low temperature regime, $\beta > \beta_c$, the equilibrium measure is concentrated on these configurations of minimal energy. But in the high temperature regime, $\beta < \beta_c$, the invariant measure does not charge too much these configurations with minimal energy. In fact the invariant measure has its mass concentrated on configurations with energy of order $-\beta N$. This follows from results in [8]. In a certain sense, when $\beta < \beta_c$, it is therefore 'un-natural' to compute the thermalization time starting from a configuration of minimal energy.
We shall therefore change our point of view: instead of considering any initial law, we shall rather estimate the time to equilibrium when the dynamics starts from the uniform probability. Doing this we expect the dynamics to avoid the configurations of minimal energy (in the high temperature regime), and thus we hope to see a dynamical phase transition.

Using generalized Poincaré inequalities, we get upper bounds for the time to equilibrium starting from the uniform law, say T_N. We prove that, when $\beta < \beta_c$, then

$$\limsup \frac{1}{N} \log T_N \leq \beta^2 \tag{4.6}$$

Comparing (4.6) with (4.5), one sees that the thermalisation time is much shorter than the inverse spectral gap. In other words, in the high temperature regime, starting from the uniform law, the dynamics reaches equilibrium much faster than starting from one of the configurations of minimal energy. These results can be interpreted as a first step towards a proof of the existence of a dynamical phase transition. Actually we expect (4.6) to be sharp i.e. we expect $\frac{1}{N}\log T_N$ to converge to β^2, for all $\beta < \beta_c$. In the low temperature regime, the asymptotics of T_N should be given by the inverse spectral gap i.e. one expects $\frac{1}{N}\log T_N$ to converge to $-\beta\beta_c$ for all $\beta \geq \beta_c$. Thus one would see the dynamical phase transition for the Metropolis dynamics.
Remember that the Hamiltonians $H(\sigma), \sigma \in \{-1,+1\}^N$ of the REM form a family of i.i.d, Gaussian random variables with mean zero and variance N, defined on some probability space, say $(\Omega, \Sigma, \mathcal{Q})$. Given $\beta \geq 0$, the inverse temperature, the Gibbs measure is defined by

$$\pi_\beta(\sigma) \equiv \frac{e^{-\beta H(\sigma)}}{Z_N} \tag{4.7}$$

where Z_N is defined in (4.2). For a given realization of the Hamiltonian, we consider the Metropolis dynamics, $X(t) = X_N(t)$: $X(t)$ is the continuous time Markov process defined on $\mathcal{X} \equiv \{-1,+1\}^N$ by the transition rates:

$$P(\sigma,\sigma') = \begin{cases} \frac{1}{N}\exp\{-\beta(H(\sigma')-H(\sigma))^+\} & \text{if } ||\sigma'-\sigma|| = 1 \\ 0 & \text{if } ||\sigma'-\sigma|| > 1 \end{cases} \tag{4.8}$$

where $a^+ = \max\{a, 0\}$ and $||x|| = \frac{1}{2}\sum_{i=1}^N |x_i|$. π_β is invariant, ergodic, and reversible for this dynamics.
The associated Dirichlet form on $L_2(\mathcal{X}, \pi_\beta)$ is given by

$$\mathcal{E}(f,g) = \frac{1}{2NZ_N(\beta)} \sum_{x,y} (f(x)-f(y))(g(x)-g(y)) e^{-\beta(H(x)\vee H(y))} \tag{4.9}$$

With the notation of part II,

$$Q(e) = \frac{1}{NZ_N(\beta)} e^{-\beta(H(x)\vee H(y))}$$

for $e=(x,y)$ with $\|x-y\|=1$.
From (4.5), one deduces that for any fixed initial law η, and any $\gamma>\beta\beta_c$, then, $\mathcal{Q}$.a.s.

$$d_{TV}(\mathcal{L}_\eta(X(e^{\gamma N})),\pi_\beta)\to 0 \tag{4.10}$$

From now on, we assume that $\beta\le\beta_c$. Given a probability measure η on $\mathcal{X}$, and $t\in I\!R$, let $\mathcal{L}_\eta(X(t))$ be the law of the process at time t starting from the initial measure η.
Given $\epsilon>0$, $c>0$ and a probability measure η, we define the time $T_N(\epsilon,c,\eta)$ to reach equilibrium starting from η, up to ϵ on a subset of $\mathcal{Q}$-probability greater than $1-e^{-cN}$ by

$$T_N(\epsilon,c,\eta)\equiv\inf\left\{T\ge 0:\mathcal{Q}\left[\sup_{s\ge T}d_{TV}\left(L_\eta(X(s)),\pi_\beta\right)\le\epsilon\right]\ge 1-e^{-cN}\right\} \tag{4.11}$$

The main result of this section is

Theorem 4.1 *Let η be the uniform probability measure on $\mathcal{X}$. Then for all $c>0$, $\epsilon>0$ and for all $\beta\le\beta_c$*

$$\limsup_{N\uparrow\infty}\frac{1}{N}\log T_N(\epsilon,c,\eta)\le\beta^2 \tag{4.12}$$

We can also prove estimates when ε goes to 0 as $N\to\infty$. We consider two cases: ε going to 0 polynomialy and as a stretched exponential.

Theorem 4.2 *Let η be the uniform probability measure on $\mathcal{X}$. There exists a constant $c_1>0$, such that for all $c>0$, there exists a constant $C_0=C_0(c,\beta)$ such that*

$$\begin{aligned}&\frac{1}{N}\log T_N(e^{-N^{1/4}(\log N)^{3/4}},c,\eta)\\ \le&\beta^2+2\beta\beta_c\left(\frac{c_1(1+c)\log N}{N}\right)^{1/2}+c_2(\beta,c)\left(\frac{\log N}{N}\right)^{1/4}+C_0(\frac{\log N}{N})^{3/4}\end{aligned} \tag{4.13}$$

where

$$c_2(\beta,c)\equiv\beta\left(12\beta\beta_c\sqrt{c_1(1+c)}\right)^{1/2}+\frac{1}{4}\frac{\beta^2+\beta_c^2}{\beta\beta_c\sqrt{c_1(1+c)}} \tag{4.14}$$

Moreover for all $\delta>0$

$$\begin{aligned}&\frac{1}{N}\log T_N(N^{-\delta},c,\eta)\le\beta^2+\beta\left(12\beta\beta_c\sqrt{c_1(1+c)}\right)^{1/2}(\frac{\log N}{N})^{1/4}\\ &+2\beta\beta_c\left(\frac{c_1(1+c)\log N}{N}\right)^{1/2}+\frac{1}{4}\frac{\beta^2+\beta_c^2}{\beta\beta_c}\frac{\delta}{\sqrt{c_1(1+c)}}(\frac{\log N}{N})^{1/2}+C_0\frac{\log N}{N}\end{aligned} \tag{4.15}$$

As a corollary, we get

Corollary 4.3 *Let η be the uniform probability measure on $\mathcal{X}$. For all $\gamma > \beta^2$, with a $\mathcal{Q}$-probability 1, for all but a finite number of indices N*

$$d_{TV}\left(\mathcal{L}_\eta(X(e^{\gamma N})), \pi_\beta)\right) \leq e^{-N^{1/4}(\log N)^{3/4}} \tag{4.16}$$

Moreover for all $\delta > 0$, if

$$t_N = \exp[\beta^2 N + \sqrt{12\beta\beta_c} N^{3/4}(\delta \log N)^{1/4} + (2\beta\beta_c + \delta(\beta^2 + \beta_c^2))(c_1 N \log N)^{1/2}] \tag{4.17}$$

then

$$d_{TV}\left(\mathcal{L}_\eta(X(t_N)), \pi_\beta)\right) \leq \frac{1}{N^\delta} \tag{4.18}$$

The error terms in the bound (4.13) have no reason to be optimal. However in (4.5) the order of magnitude of the error terms are optimal as it was observed in [4].

To prove the theorems we will need estimates for the constants, $\mathcal{L}_{\pi_\beta}(p)$ and $\mathcal{L}_\eta(p)$ using (2.19). This will be done now and the result will be collected in the Proposition 4.9.

These estimates will also depend on the choice of paths $\gamma(x, y)$. To estimate the spectral gap, the following set of paths was introduced in [4] and they work also here: given $i \in \{1, \ldots, N\}$, and $x, y \in \mathcal{X}$, such that $x_i \neq y_i$ let $\gamma^i(x, y)$ be the path starting at x and ending at y obtained by flipping the disagreeing spins, starting at the site i and then going cyclically. Let $\Gamma^i = \{\gamma^i(x, y), x, y \in \mathcal{X}\}$. Given x, y and $\gamma(x, y)$, let $\overline{\gamma(x, y)}$ be the set of points visited by the path and $\gamma^o(x, y) = \overline{\gamma(x, y)} \setminus \{x, y\}$ the set of the interior points of the path. Note that if the number of discrepancies between x and y is n then there exist n interior disjoint paths in $\{\gamma^i(x, y), i = 1, \ldots, N\}$. This comes from the fact that if $i_1, \ldots, i_n$ are the n sites where x and y disagree, then the paths $\gamma^{i_1}(x, y), \ldots, \gamma^{i_n}(x, y)$ are interior disjoint. The set of paths we will construct will depend on the realization of $H(x)$: it is a random set. Given a positive number c_e, we will say that a point z is good if $H(z) \leq \sqrt{(1 + c_e)2N \log N}$. Call G the set of good points. If z is not good, we call it 'bad' and write B for the set of bad points. A path is good if all its interior points are good. Note that we need to select a path for any pair of points (x, y), and the typical number of bad points is of order $2^N 2^{-(1+c_e)\log N}$. We cannot neglect good paths $\gamma(x, y)$ with bad end points x or y or both. We construct the set of paths Γ according to the following rules:
For $||x - y||_1 \geq \frac{N}{\log N}$, if there is a good path in $\{\gamma^i(x, y), i = 1, \ldots, N\}$, choose the first and put it in Γ ; otherwise, choose $\gamma^1(x, y)$. For $||x - y||_1 < \frac{N}{\log N}$, if there exists a good site z in $\mathcal{X}$ such that $||x - z|| \geq \frac{N}{\log N}$, $||y - z|| \geq \frac{N}{\log N}$ and if there are good paths, one in $\{\gamma^i(x, z), i = 1, \ldots, N\}$ and another in $\{\gamma^i(z, y), i = 1, \ldots, N\}$ such that the union of these two good paths is a self avoiding path, then we select this union as the path connecting x

and y in Γ. (Note that this is a good path since z is good); otherwise, select $\gamma^1(x,y)$. Note that all the paths constructed in this way have length smaller than N. A fundamental result that can be easily proven by keeping the $\mathcal{Q}$-probability in the proof of the proposition 4 .1 in [4] is

Proposition 4.4 *For all $c_e > 0$, there exists $N_0(c_e)$ such that for all $N \geq N_0(c_e)$, with a $\mathcal{Q}$-probability $\geq 1 - e^{-c_e N}$, all the paths of the previous set Γ are good i.e. satisfy $H(z) \leq \sqrt{(1+c_e)N \log N}$, for all $z \in \gamma(x,y) \setminus (x,y)$, for all $(x,y) \in \mathcal{X}^2$. Moreover they have a length smaller than N.*

We say that an edge $e = (x, x')$ is good, if x and y are good, this will be denoted by $e \in \mathcal{G}$, otherwise the edge is bad: $e \in \mathcal{B}$. Note the important fact that, with our construction, a given edge $e = (x, x')$ belonging to Γ can have at most one bad point among x and x'.

Let us first estimate, $\mathcal{L}_{\pi_\beta}(p)$, see (2.19). The weights $\lambda(x)$ are chosen in the following way: Let d be such that $\beta < d \leq \beta_c$, to be chosen later. We set $d = \beta(1+\zeta)$ with $0 < \zeta < (\beta_c - \beta)/\beta$. Let

$$\lambda(x) = \begin{cases} 1 & \text{if } H(x) \geq -dN \\ \lambda & \text{otherwise.} \end{cases} \tag{4.19}$$

where

$$\begin{aligned} \lambda &\equiv \left(\sum_{x \in \mathcal{X}} e^{-\beta H(x)} \mathbb{I}_{\{H(x) \leq -dN\}} \right)^\rho \\ &\equiv (Z_N(\beta, \leq -d))^\rho \end{aligned} \tag{4.20}$$

for some $\rho > 0$ to be chosen later.
First we consider the first term in the right hand side of (2.19) the other ones will be treated later. Let us denote

$$R(d, \rho, p) \equiv \sum_{x \in \mathcal{X}} \pi_\beta(x) (\lambda(x))^{p/1-p} \tag{4.21}$$

Lemma 4.5 *Let $\zeta > 0$, $0 < \rho < 1$ and $0 < p < 1/2$ that satisfy $0 < \zeta \leq (\beta_c - \beta)/\beta$ and*

$$\frac{p\rho}{\zeta^2(1-p)} \leq \frac{1}{2} \frac{\beta^2}{\beta^2 + \beta_c^2} \tag{4.22}$$

There exists an absolute constant c_1, and, for any $c > 0$, there exists $N_0(\beta, c, \zeta)$ such that for any $N \geq N_0(\beta, c, \zeta)$ such that

$$\sqrt{\frac{N}{(1+c)\log N}} \geq \frac{12\beta_c}{\zeta^2 \beta} \sqrt{c_1} \tag{4.23}$$

then, with a $\mathcal{Q}$-probability$\geq 1-e^{-cN}$, we have

$$R(d,\rho,p) \leq 2 \tag{4.24}$$

where $d=\beta(1+\zeta)$

Proof: Let us denote

$$R^1(d,\rho,p) \equiv \sum_{x\in\mathcal{X}} \pi_\beta(x)\mathbb{I}_{\{H(x)\geq -dN\}} \tag{4.25}$$

and $R(d,\rho,p) \equiv R^1(d,\rho,p) + R^2(d,\rho,p)$. We have $R^1(d,\rho,p) \leq 1$ and, to estimate $R^2(d,\rho,p)$, we use the following lemma that will be proved in the section VI.

Lemma 4.6 *There exits a constant c_1, such that for all $c>0$, there exists a $N_0(\beta,c,\zeta)$ such that for all $N \geq N_0(\beta,c,\zeta)$ with a $\mathcal{Q}$-probability $\geq 1-e^{-cN}$*

$$Z_N(\beta,\leq -d) \leq e^{\beta\beta_c\sqrt{c_1(1+c)N\log N}} e^{N[\beta d - \frac{d^2}{2}+\frac{\beta_c^2}{2}]} \tag{4.26}$$

and

$$Z_N(\beta) \geq e^{-\beta\beta_c\sqrt{c_1(1+c)N\log N}} e^{N[\frac{\beta^2}{2}+\frac{\beta_c^2}{2}]} \tag{4.27}$$

Note that

$$R^2(d,\rho,p) = \frac{Z_N(\beta,\leq -d)^{1+\rho p/(1-p)}}{Z_N(\beta)}$$

Therefore Lemma 4.6 implies that

$$R^2(d,\rho,p) \leq e^{[2+\frac{p\rho}{1-p}]\beta\beta_c\sqrt{c_1(1+c)N\log N}} e^{-N[\frac{(d-\beta)^2}{2}]} e^{N\frac{p\rho}{1-p}[\beta d-\frac{d^2}{2}+\frac{\beta_c^2}{2}]} \tag{4.28}$$

Now using (4.22), we get

$$\frac{\rho p}{1-p}\left[\beta d - \frac{d^2}{2}+\frac{\beta_c^2}{2}\right] \leq \frac{1}{2}\frac{(d-\beta)^2}{2} \tag{4.29}$$

Using $0<\rho<1$ and $0<p<1/2$, we have $\rho p/(1-p) \leq 1$ therefore (4.23) implies that

$$\left[2+\frac{p\rho}{1-p}\right]\beta\beta_c\sqrt{c_1(1+c)N\log N} \leq \frac{1}{2}\frac{(d-\beta)^2}{2}N \tag{4.30}$$

from which we immediately get (4.24). ∎

Now we estimate the other term in the right hand side of (2.19). Let us denote

$$\frac{1}{2\mathcal{L}^*_{\pi_\beta}(1)} \equiv \sum_e \frac{1}{Q(e)} \left[\sum_{x,y:\gamma(x,y)\ni e} \frac{\pi_\beta(x)}{\lambda(x)} \frac{\pi_\beta(y)}{\lambda(y)} \right]^2 \tag{4.31}$$

Proposition 4.7 *We assume that $2(1-\rho) < 1$.*
There exists a constant c_1, such that for all $c > 0$ and all $\zeta > 0$, $\zeta < (\beta_c - \beta)/\beta$, satisfying (4.22), there exist $N_0(\beta, c, \zeta)$ such that for all $N \geq N_0(\beta, c)$, with a $\mathcal{Q}$-probability $\geq 1 - e^{-cN}$, we have

$$\frac{1}{2\mathcal{L}^*_{\pi_\beta}(1)} \leq 22N^4 e^{2\beta\beta_c\sqrt{c_1(1+c)N\log N}} e^{\beta dN} \tag{4.32}$$

Proof: Let us write

$$\frac{1}{2\mathcal{L}^*_{\pi_\beta}(1)} = L_{\mathcal{B}} + L_{\mathcal{G}} \tag{4.33}$$

where $L_{\mathcal{G}}$ is the same as (4.31) but with the sum $\sum_e$ restricted to good edges and $L_{\mathcal{B}}$ with bad edges.
Let us first consider $L_{\mathcal{B}}$. Using convexity and symmetry,we can write

$$L_{\mathcal{B}} \leq 3L_{\mathcal{B}}(\geq, \geq) + 6L_{\mathcal{B}}(\geq, <) \tag{4.34}$$

where $L_{\mathcal{B}}(\geq, \geq)$ is the same as in (4.31) but with the following restrictions: $e \in \mathcal{B}, H(x) \geq -dN, H(y) \geq -dN$. $L_{\mathcal{B}}(\geq, <)$ is defined similarly.
Let $U \equiv \{x\ ;\ H(x) \geq -dN\}$ and $D \equiv \{x\ ;\ H(x) < -dN\}$. Since a bad edge is the first or the last edge of the path, if $e = (z, z') \in \gamma(x, y)$ is a bad edge then we have either $z \in B$ and $x = z$ or $z' \in B$ and $z' = y$. By symmetry it is sufficient to consider the first case. Then $1/Q(e) = NZ_N(\beta)\exp(\beta H(z))$. Note in particular that it is not possible to have $e = (z, z') \in \gamma(x, y)$, e bad and both x and y belonging to D. This is the reason why we do not have a term $L_{\mathcal{B}}(<, <)$.

$$\begin{aligned} L_{\mathcal{B}}(\geq, \geq) &\leq \frac{2N}{Z_N^3(\beta)} \sum_{e=(z,z')} e^{\beta H(z)} \left[\sum_{x,y \in U\times U} e^{-\beta H(x)} e^{-\beta H(y)} \mathbb{1}_{\{x=z\}} \right]^2 \\ &= \frac{2N}{Z_N^3(\beta)} \sum_{e=(z,z')} e^{-\beta H(z)} \left[\sum_{y\in U} e^{-\beta H(y)} \right]^2 \leq 2N \end{aligned} \tag{4.35}$$

Using similar arguments and recalling (4.20), we get

$$
\begin{aligned}
L_{\mathcal{B}}(\geq,<) &\leq \frac{2N}{Z_N^3(\beta)} \sum_{e=(z,z')} e^{\beta H(z)} \left[\sum_{x,y\in U\times D} e^{-\beta H(x)} \frac{e^{-\beta H(y)}}{Z_N^{\rho}(\beta,\leq -d)} 1\!\!1_{\{x=z\}} \right]^2 \\
&\leq \frac{2N}{Z_N^3(\beta)} \sum_{e=(z,z')} e^{-\beta H(z)} \left[\sum_{y\in D} \frac{e^{-\beta H(y)}}{Z_N^{\rho}(\beta,\leq -d)} \right]^2 \\
&\leq \frac{2N}{Z_N^2(\beta)} \left[Z_N(\beta,\leq -d) \right]^{2(1-\rho)}
\end{aligned}
\tag{4.36}
$$

Using (4.26), (4.27), and $2(1-\rho)\leq 1$ we get

$$
\begin{aligned}
L_{\mathcal{B}}(\geq,<) &\leq 2N e^{3\beta\beta_c\sqrt{c_1(1+c)N\log N}} e^{-N[\frac{\beta^2+\beta_c^2+(d-\beta)^2}{2}]} \\
&\leq 2N e^{-N[\frac{\beta^2+\beta_c^2}{2}]} \\
&\leq 2N
\end{aligned}
\tag{4.37}
$$

where we have used (4.23) at the second step. We have proved that

$$
L_{\mathcal{B}} \leq 18N \tag{4.38}
$$

We consider now $L_{\mathcal{G}}$ As before, using convexity and symmetry, we write

$$
L_{\mathcal{G}} \leq 4L_{\mathcal{G}}(\geq,\geq) + 8L_{\mathcal{G}}(\geq,<) + 4L_{\mathcal{G}}(<,<) \tag{4.39}
$$

We first consider $L_{\mathcal{G}}(<,<)$. Since for a good edge $e=(z,z')$, we have $H(z)\vee H(z')\leq \sqrt{(1+c_e)N\log N}$, we therefore get

$$
L_{\mathcal{G}}(<,<) \leq \frac{N e^{\beta\sqrt{(1+c_e)N\log N}}}{Z_N^3(\beta)} \sum_{e\in\mathcal{G}} \left[\sum_{x,y\in D\times D} 1\!\!1_{\{\gamma(x,y)\ni e\}} \frac{e^{-\beta H(x)}}{Z_N^{\rho}(\beta,\leq -d)} \frac{e^{-\beta H(y)}}{Z_N^{\rho}(\beta,\leq -d)} \right]^2 \tag{4.40}
$$

On one hand we have

$$
\left[\sum_{x,y\in D\times D} 1\!\!1_{\{\gamma(x,y)\ni e\}} \frac{e^{-\beta H(x)}}{Z_N^{\rho}(\beta,\leq -d)} \frac{e^{-\beta H(y)}}{Z_N^{\rho}(\beta,\leq -d)} \right] \leq Z_N^{2(1-\rho)}(\beta,\leq -d) \tag{4.41}
$$

On the other hand we have

$$
\begin{aligned}
&\sum_{e\in\mathcal{G}}\sum_{x,y\in D\times D} \mathbb{1}_{\{\gamma(x,y)\ni e\}} \frac{e^{-\beta H(x)}}{Z_N^{\rho}(\beta,\leq -d)}\frac{e^{-\beta H(y)}}{Z_N^{\rho}(\beta,\leq -d)} \\
&= \sum_{x,y\in D\times D} \frac{e^{-\beta H(x)}}{Z_N^{\rho}(\beta,\leq -d)}\frac{e^{-\beta H(y)}}{Z_N^{\rho}(\beta,\leq -d)} \sum_{e\in\mathcal{G}} \mathbb{1}_{\{\gamma(x,y)\ni e\}} \\
&\leq N Z_N^{2(1-\rho)}(\beta,\leq -d)
\end{aligned}
\tag{4.42}
$$

where at the last step we have used that the length of a path is smaller than N. Therefore using (4.41) and (4.42) in (4.40), then (4.26) and (4.27) and at last $4(1-\rho)\leq 2\leq 3$ we get

$$
\begin{aligned}
L_{\mathcal{G}}(<,<) &\leq N^2 e^{\beta\sqrt{(1+c_e)N\log N}} \frac{Z_N^{4(1-\rho)}(\beta,\leq -d)}{Z_N^3(\beta)} \\
&\leq N^2 e^{\beta\sqrt{(1+c_e)N\log N}} e^{6\beta\beta_c\sqrt{c_1(1+c)N\log N}} e^{-N\frac{3(d-\beta)^2}{2}} \\
&\leq N^2 e^{7\beta\beta_c\sqrt{c_1(1+c)N\log N}} e^{-N\frac{3(d-\beta)^2}{2}} \leq N^2
\end{aligned}
\tag{4.43}
$$

Where we have used $\beta_c=\sqrt{2\log 2}>1$, (4.23) and we have chosen $c_e=c$ and $c_1>1$.

Consider now $L_{\mathcal{G}}(\geq,<)$. Using exactly the same kind of arguments, using (4.26) and (4.27), and $2(1-\rho)\leq 1$ we get

$$
\begin{aligned}
L_{\mathcal{G}}(\geq,<) &\leq N^2 e^{\beta\sqrt{c_1(1+c)N\log N}} \frac{Z_N^{2(1-\rho)}(\delta,\leq -d)}{Z_N(\beta)} \\
&\leq N^2 e^{3\beta\beta_c\sqrt{c_1(1+c)N\log N}} e^{-N\frac{(d-\beta)^2}{2}} \leq N^2
\end{aligned}
\tag{4.44}
$$

where at the last step we have used (4.23).

We consider now $L_{\mathcal{G}}(\geq,\geq)$. Since the edge is good, we have

$$
\begin{aligned}
L_{\mathcal{G}}(\geq,\geq) &\leq \frac{Ne^{\beta\sqrt{(1+c_e)N\log N}}}{Z_N^3(\beta)}\sum_{e\in\mathcal{G}}\left[\sum_{x,y\in U\times U}\mathbb{1}_{\{\gamma(x,y)\ni e\}}e^{-\beta H(x)}e^{-\beta H(y)}\right]^2 \\
&\leq Ne^{\beta\sqrt{(1+c_e)N\log N}}\sup_e\left[\sum_{x,y\in U\times U}\mathbb{1}_{\{\gamma(x,y)\ni e\}}\frac{e^{-\beta H(x)}e^{-\beta H(y)}}{Z_N(\beta)}\right] \\
&\quad\times\sum_{e\in\mathcal{G}}\left[\sum_{x,y\in U\times U}\mathbb{1}_{\{\gamma(x,y)\ni e\}}\frac{e^{-\beta H(x)}e^{-\beta H(y)}}{Z_N^2(\beta)}\right] \\
&\leq N^2e^{\beta\sqrt{(1+c_e)N\log N}}\sup_e\left[\sum_{x,y\in U\times U}\mathbb{1}_{\{\gamma(x,y)\ni e\}}\frac{e^{-\beta H(x)}e^{-\beta H(y)}}{Z_N(\beta)}\right]
\end{aligned}
\tag{4.45}
$$

To continue we will need an adaptation of [4]. Let us call

$$
\Lambda(d)\equiv\sup_e\left[\sum_{x,y\in U\times U}\mathbb{1}_{\{\gamma(x,y)\ni e\}}\frac{e^{-\beta H(x)}e^{-\beta H(y)}}{Z_N(\beta)}\right]
\tag{4.46}
$$

Recalling that the paths in Γ are constructed using paths in $\cup_{i=1}^N\Gamma^i$, we get immediately

$$
\Lambda(d)\leq N\sup_{1\leq i\leq N}\Lambda^{(i)}(d)
\tag{4.47}
$$

where $\Lambda^{(i)}(d)$ is as in (4.46) but with paths in Γ^i. It is enough to consider the case $i=1$ the other ones being similar. Now for a given edge $e=(z,z')$, there exists a $j\in\{1,\dots,N\}$ such that $z'=z^j$, that is z' is the configuration obtained from z by flipping the spin at the site j. Note at this point that the set of all $(x,y):\gamma(x,y)\ni e$ for $\gamma\in\Gamma^1$ is exactly

$$
\bigcup_{x\in\{-1,+1\}^{j-1}}\ \bigcup_{y\in\{-1,+1\}^{N-j}}((x_1,\dots,x_{j-1},z_j,\dots,z_N),(z_1,\dots,z_{j-1},-z_j,y_{j+1},\dots,y_N))
\tag{4.48}
$$

Denoting $z_{>j}\equiv(z_{j+1},\dots,z_N)$, $z_{<j}\equiv(z_1,\dots,z_{j-1})$,

$$
Z_{j-1}^{(1)}(\beta,\geq -d)[z_j,z_{>j}]\equiv\sum_{x\in\{-1,+1\}^{j-1}}e^{-\beta H(x,z_j,z_{>j})}\mathbb{1}_{\{H(x,z_j,z_{>j})\geq -dN\}}
\tag{4.49}
$$

and

$$
Z_{N-j}^{(1)}(\beta,\geq -d)[z_{<j},-z_j]\equiv\sum_{y\in\{-1,+1\}^{N-j}}e^{-\beta H(z_{<j},-z_j,y)}\mathbb{1}_{\{H(z_{<j},-z_j,y)\geq -dN\}}
\tag{4.50}
$$

we get immediately:

$$\Lambda^{(1)}(d) \leq \sup_{z\in\mathcal{X}} \frac{1}{Z_N(\beta)} Z^{(1)}_{j-1}(\beta, \geq -d)[z_j, z_{>j}] Z^{(1)}_{N-j}(\beta, \geq -d)[z_{<j}, -z_j] \tag{4.51}$$

To continue, we need the following lemma that will be proved in the next section.

Lemma 4.8 *There exists a constant $c_1 > 0$, such that for all $c > 0$, if $c_u = c_1^{-1}(2\log 2 + c)$, then we can find an $N_0 = N_0(\beta, c)$ such that for all $N > N_0(\beta, c)$, with a $\mathcal{Q}$-probability $\geq 1 - e^{-cN}$, if we call $M \equiv \sqrt{N/\log_2(Nc_u)}$ and $j - 1 \equiv \alpha N$ then*

$$\sup_{z_j, z_{>j}} Z^{(1)}_{j-1}(\beta, \geq -d)[z_j, z_{>j}] \leq \sqrt{N} e^{\beta\beta_c \sqrt{N \log_2(c_u N)}} \left(2^j + \mathcal{Z}_N(\beta, d, \alpha)\right) \tag{4.52}$$

where

$$\mathcal{Z}_N(\beta, d, \alpha) = \begin{cases} e^{\beta d N} & \text{if } \sqrt{\alpha M^2 - 1} < \frac{\beta}{\beta_c} M \\ e^{\beta d N} + e^{N[\frac{\beta^2}{2} + \alpha \frac{\beta_c^2}{2}]} & \text{if } \frac{\beta}{\beta_c} M \leq \sqrt{\alpha M^2 - 1} < \frac{d}{\beta_c} M \\ e^{N[\frac{\beta^2}{2} + \alpha \frac{\beta_c^2}{2}]} & \text{if } \frac{d}{\beta_c} M \leq \sqrt{\alpha M^2 - 1} \end{cases} \tag{4.53}$$

Now inserting (4.53) for $j - 1 = \alpha N$ and $N - j = (1-\alpha)N$ in (4.51), considering the nine resulting terms, using

$$\beta d < \beta\beta_c < \frac{\beta^2}{2} + \frac{\beta_c^2}{2} \tag{4.54}$$

to simplify the computations, and maximizing over $\alpha \in [0, 1]$, it is just a long task to get

$$\begin{aligned} \Lambda^{(1)}(d) &\leq \sqrt{N} e^{\beta\beta_c \sqrt{N \log_2(c_u N)}} e^{\beta d N} \\ &\leq \sqrt{N} e^{\beta\beta_c \sqrt{c_1(1+c) N \log N}} e^{\beta d N} \end{aligned} \tag{4.55}$$

with a $\mathcal{Q}$-probability $\geq 1 - e^{-cN}$. Inserting (4.55) and (4.47) in (4.45), we get

$$L_{\mathcal{G}}(\geq, \geq) \leq N^4 e^{2\beta\beta_c \sqrt{c_1(1+c) N \log N}} e^{\beta d N} \tag{4.56}$$

Using (4.33),(4.38),(4.43),(4.44) and (4.56) we get (4.32). ■

Now we estimate $\mathcal{L}_\eta(p)$ see (2.19), when η is the uniform measure on $\mathcal{X}$. We take the weights $\mu(x) = 1$. Since we have already estimated the first factor in Lemma 4.5, it remains to estimate

$$\frac{1}{\mathcal{L}_\eta^*(1)} \equiv \sum_e \frac{1}{Q(e)} \left[\sum_{x,y:\gamma(x,y)\ni e} \frac{\pi_\beta(x)}{\lambda(x)} \frac{1}{2^N} \right]^2 \tag{4.57}$$

Proposition 4.9 *There exists a constant c_1, such that for all $c > 0$, for all $\zeta > 0$, $\zeta < (\beta_c - \beta)/\beta$, there exists $N_0(\beta, c, \zeta)$, such that for all N that satisfy (4.23) and are larger than $N_0(\beta, c, \zeta)$, with a $\mathcal{Q}$-probability $\geq 1 - e^{-cN}$,*

$$\frac{1}{2\mathcal{L}_\eta^*(1)} \leq 4N^2 e^{4\beta\beta_c\sqrt{c_1(1+c)N\log N}} e^{\beta dN} \tag{4.58}$$

Proof: As before by considering separately the cases where $e \in \mathcal{G}$ and $e \in \mathcal{B}$, we write

$$\frac{1}{2\mathcal{L}_\eta^*(1)} \equiv L_{\mathcal{B}}(\eta) + L_{\mathcal{G}}(\eta) \tag{4.59}$$

Distinguishing bad and good edges and separating the cases $x \in D$ or $x \in U$, we get four terms that we call, $L_{\mathcal{B}}(\eta, \geq)$, $L_{\mathcal{B}}(\eta, <)$, $L_{\mathcal{G}}(\eta, \geq)$ and $L_{\mathcal{G}}(\eta, <)$.
Let us start with $L_{\mathcal{B}}(\eta, <)$. We should then have $y = z' \in B$. Therefore

$$\begin{aligned} L_{\mathcal{B}}(\eta, <) &\leq \frac{2N}{Z_N(\beta)} \sum_{e=(z,z')} e^{\beta H(z')} \left[\sum_{x\in D,y} \frac{\pi_\beta(x)}{Z_N^\rho(\beta, \leq -d)} \frac{1}{2^N} \mathbb{I}_{\{y=z'\}} \right]^2 \\ &\leq \frac{2N}{Z_N(\beta)} \sum_{z'} e^{\beta H(z')} \left[\frac{Z_N^{(1-\rho)}(\beta, \leq -d)}{2^N} \right]^2 \\ &= \frac{2N}{Z_N(\beta)} Z_N(-\beta) \frac{Z_N^{2(1-\rho)}(\beta, \leq -d)}{2^{2N}} \end{aligned} \tag{4.60}$$

Now since it is clear that $Z_N(\beta)$ and $Z_N(-\beta)$ have the same distribution and therefore satisfy the same estimates, we get that with a $\mathcal{Q}$-probability $\geq 1 - e^{-cN}$,

$$\frac{Z_N(-\beta)}{Z_N(\beta)} \leq e^{2\beta\beta_c\sqrt{c_1(1+c)N\log N}} \tag{4.61}$$

Using now (4.26), and $2(1-\rho) < 1$ we get

$$L_{\mathcal{B}}(\eta, <) \leq 2Ne^{3\beta\beta_c\sqrt{c_1(1+c)N\log N}} e^{N[2\beta d - d^2]} \tag{4.62}$$

Consider now $L_{\mathcal{B}}(\eta, \geq)$. Then e is bad and $x \in U$. We have to deal separately with, case 1, $x = z \in B$and, case 2, $y = z' \in B$. By convexity

$$L_{\mathcal{B}}(\eta, \geq) \leq 2L_{\mathcal{B}}(\eta, \geq, 1) + 2L_{\mathcal{B}}(\eta, \geq, 2) \tag{4.63}$$

On the one hand we have

$$\begin{aligned} L_{\mathcal{B}}(\eta, \geq, 1) &\leq \frac{N}{Z_N(\beta)} \sum_z e^{-\beta H(z)} \left[\sum_y \frac{1}{2^N}\right]^2 \\ &= N \end{aligned} \tag{4.64}$$

On the other hand we have

$$\begin{aligned} L_{\mathcal{B}}(\eta, \geq, 2) &\leq \frac{N}{Z_N(\beta)} \sum_{z'} e^{-\beta H(z')} \left[\sum_{x \in U} e^{-\beta H(x)} \frac{1}{2^N}\right]^2 \\ &\leq \frac{N}{Z_N(\beta)} Z_N(-\beta) \left[\frac{Z_N(\beta)}{2^N}\right]^2 \\ &\leq N e^{4\beta\beta_c \sqrt{c_1(1+c)N \log N}} e^{\beta^2 N} \end{aligned} \tag{4.65}$$

Collecting (4.62), (4.64) and (4.65), we get

$$\begin{aligned} L_{\mathcal{B}}(\eta) &\leq 2N e^{4\beta\beta_c \sqrt{c_1(1+c)N \log N}} \left(e^{\beta^2 N} + e^{[2\beta d - d^2]N}\right) \\ &\leq 4N e^{4\beta\beta_c \sqrt{c_1(1+c)N \log N}} e^{\beta d N} \end{aligned} \tag{4.66}$$

where, at the last step, we used that $d \geq \beta$ and therefore $2\beta d - d^2 \leq \beta d$.
Consider now $L_{\mathcal{G}}(\eta, <)$. Since we consider now good edges, we have

$$\begin{aligned} L_{\mathcal{G}}(\eta, <) &\leq \frac{N e^{\beta\sqrt{(1+c_e)N \log N}}}{Z_N(\beta)} \sum_{e \in \mathcal{G}} \left[\sum_{x \in D, y} \frac{e^{-\beta H(x)}}{Z_N^{\rho}(\beta, \leq -d)} \frac{1}{2^N}\right]^2 \\ &\leq \frac{N e^{\beta\sqrt{(1+c_e)N \log N}}}{Z_N(\beta)} Z_N^{(1-\rho)}(\beta \leq -d) \sum_{e \in \mathcal{G}} \left[\sum_{x \in D, y} \frac{e^{-\beta H(x)}}{Z_N^{\rho}(\beta, \leq -d)} \frac{1}{2^N}\right] \\ &\leq \frac{N^2 e^{\beta\sqrt{(1+c_e)N \log N}}}{Z_N(\beta)} Z_N^{2(1-\rho)}(\beta \leq -d) \\ &\leq N^2 e^{3\beta\beta_c \sqrt{c_1(1+c)N \log N}} e^{-\frac{1}{2}(d-\beta)^2 N} \\ &\leq N^2 \end{aligned} \tag{4.67}$$

where we used that $2(1-\rho) < 1$ and (4.23) at the last step.
It remains to consider $L_{\mathcal{G}}(\eta, \geq)$. Using the fact that e is good, we get

$$
\begin{aligned}
L_{\mathcal{G}}(\eta, \geq) &\leq \frac{N e^{\beta\sqrt{(1+c_e)N\log N}}}{Z_N(\beta)} \sum_{e\in\mathcal{G}} \left[\sum_{x\in U, y, \gamma(x,y)\ni e} e^{-\beta H(x)} \frac{1}{2^N} \right]^2 \\
&\leq N e^{\beta\sqrt{(1+c_e)N\log N}} \sup_{e\in\mathcal{G}} \left[\sum_{x\in U, y, \gamma(x,y)\ni e} e^{-\beta H(x)} \frac{1}{2^N} \right] \\
&\quad \times \frac{1}{Z_N(\beta)} \sum_{e\in\mathcal{G}} \left[\sum_{x\in U, y, \gamma(x,y)\ni e} e^{-\beta H(x)} \frac{1}{2^N} \right] \\
&\leq N^2 e^{\beta\sqrt{(1+c_e)N\log N}} \sup_{e\in\mathcal{G}} \left[\sum_{x\in U, y, \gamma(x,y)\ni e} e^{-\beta H(x)} \frac{1}{2^N} \right]
\end{aligned}
\tag{4.68}
$$

To estimate this last supremum, we use a similar argument as the one we used to treat (4.51). Using (4.52), and the same notation as in (4.49), after a not too long computation, we get that with a $\mathcal{Q}$-probability $\geq 1 - e^{-cN}$

$$
\begin{aligned}
\Lambda(\eta, d) \equiv \sup_{e\in\mathcal{G}} &\left[\sum_{x\in U, y, \gamma(x,y)\ni e} e^{-\beta H(x)} \frac{1}{2^N} \right] \\
&\leq \sup_{1\leq j\leq N} \sup_{z\in\mathcal{X}} Z_{j-1}(\beta, \geq -d)[z_j, z_{>j}] 2^{-j} \\
&\leq e^{\beta\beta_c\sqrt{N\log_2(c_u N)}} e^{\beta d N}
\end{aligned}
\tag{4.69}
$$

Collecting (4.68) and (4.69), we get

$$
L_{\mathcal{G}}(\eta, \geq) \leq N^2 e^{2\beta\beta_c\sqrt{c_1(1+c)N\log N}} e^{\beta d N}
\tag{4.70}
$$

Collecting (4.66), (4.67) and (4.70), this entails (4.58). ■

Now we put together all the results concerning the quantities $\mathcal{L}_\eta(p)$ and $\mathcal{L}_{\pi_\beta}$. That is collecting Lemmata 4.5 and Proposition4.7 and 4.9, recalling (2.19) we have

Proposition 4.10 *Let $\beta < \beta_c$, $0 < \zeta < (\beta_c - \beta)/\beta$ and $0 < p < 1/2$ satisfy*

$$
\frac{p}{\zeta^2(1-p)} < \frac{\beta^2}{\beta^2 + \beta_c^2}
$$

There exists an absolute constant c_1, such that for all $c > 0$, there exists a $N_0(\beta, c, \zeta)$ such that for all $N \geq N_0(\beta, c, \zeta)$ and N satisfying condition (4.23) then, with a $\mathcal{Q}$–probability $\geq 1 - e^{-cN}$, we have

$$\frac{1}{\mathcal{L}_{\pi_\beta}(p)} \leq 4^{\frac{1-3p+p^2}{p(1-p)}} 22N^4 e^{2\beta\beta_c\sqrt{c_1(1+c)N\log N}} e^{\beta^2(1+\zeta)N} \tag{4.71}$$

and

$$\frac{1}{\mathcal{L}_{\eta}(p)} \leq 4^{\frac{2-3p+2p^2}{p(1-p)}} 4N^2 e^{4\beta\beta_c\sqrt{c_1(1+c)N\log N}} e^{\beta^2(1+\zeta)N} \tag{4.72}$$

Remark: the aim of this remark is to discuss the implications of Proposition 4.10 as far as the behaviour of the eigenvectors of the Metropolis dynamics are concerned. To simplify things, we only consider the almost sure asymptotics of the first non trivial eigenvector: assume that we have constructed the Hamiltonians $H(\sigma)$ corresponding to the different values of N on the same probability space, and fix one realisation. From Proposition 4.10, we then know that,

$$\limsup \frac{1}{N} \log \frac{1}{\mathcal{L}_{\pi_\beta}(p)} \leq \beta^2(1+\zeta) \tag{4.73}$$

Let now λ denote the spectral gap of $\mathcal{E}$. λ depends on the realisation of H and on N. And let ψ be the corresponding eigenvector. We assume that $\pi_\beta(\psi^2) = 1$. From [4], we then know that

$$\lim \frac{1}{N} \log \frac{1}{\lambda} = \beta\beta_c \tag{4.74}$$

Therefore, provided we choose ζ small enough, we will have

$$\limsup \frac{1}{N} \log \frac{\lambda}{\mathcal{L}_{\pi_\beta}(p)} \leq -a \tag{4.75}$$

, where $a > 0$ is a deterministic constant that depends on β. It then follows from (2.18) that $\pi_\beta(|\psi|) \leq \exp(-aN)$ for large enough N and with possibly a different value for the constant a. In other words the eigenvector ψ becomes concentrated on its support. As a matter of fact, this is only another way to understand the fact that thermalisation times depend a lot on the initial law: eigenvectors corresponding to low eigenvalues become singular.

Proof of Theorem:4.1 recalling (2.10), (4.24), (4.32) and (4.58), we get

$$\begin{aligned} \frac{1}{N}\log T_N(\epsilon, c, \eta) \leq \frac{1}{N}\log C_p + \frac{2-p}{pN}\log\frac{1}{\epsilon} + \frac{4\log N}{N} \\ + 2\beta\beta_c\left(c_1(1+c)\frac{\log N}{N}\right)^{1/2} + \beta^2(1+\zeta) \end{aligned} \tag{4.76}$$

where C_p is the constant in (2.10). (Remember that $d = \beta(1+\zeta)$). Now taking first the limit $N \uparrow \infty$, we get

$$\limsup \frac{1}{N} \log T_N(\epsilon, c, \eta) \leq \beta^2(1+\zeta) \tag{4.77}$$

(4.77) is satisfied for all $\zeta > 0$. (Just choose p small enough so that (4.22) is satisfied). Therefore

$$\limsup \frac{1}{N} \log T_N(\epsilon, c, \eta) \leq \beta^2$$

■

Proof of Theorem:4.2
The proof is a little more involved than the previous one. Choose

$$\log \frac{1}{\epsilon} = N^{1/4}(\log N)^{3/4}$$

$$\zeta^2 = 12\frac{\beta_c}{\beta}(c_1(1+c)\frac{\log N}{N})^{1/2}$$

$\rho = 3/4$ and

$$\frac{p}{1-p} = \frac{2}{3}\frac{\beta^2}{\beta_c^2+\beta^2}\zeta^2$$

Then (4.22) and (4.23) are satisfied. Also

$$\frac{2}{p} = \frac{1}{4}\frac{\beta^2+\beta_c^2}{\beta\beta_c}\frac{1}{\sqrt{c_1(1+c)}}\frac{1}{\sqrt{N\log N}}$$

and we deduce the upper bound (4.13) from (4.76). The proof of (4.15) is similar, with now $\log(1/\epsilon) = \delta \log N$. ■

V. The Medium from the point of view of the process

In this section, we shall consider the process of the environment as seen from the particle. This process will be denoted by ω_t. For any fixed N, let $S_N \equiv \{-1,+1\}^N$. We endow S_N with its natural group structure i.e. for $\sigma, \sigma' \in S_N$, we let $\sigma.\sigma' \in S_N$ be the configuration $(\sigma.\sigma')_i = \sigma_i\sigma'_i$. Let $\mathbb{1}$ be the configuration $(\mathbb{1})_i = 1$ for all i. For $1 \leq i \leq N$, we also define $\underline{i}$ to be the configuration whose i-th coordinate is -1, and the other coordinates are $+1$. Thus $\sigma.\underline{i}$ is the configuration obtained by flipping the i-th coordinate of σ.

Without loss of generality, we may, and will assume that our random Hamiltonian H is defined on the canonical space $\Omega \equiv I\!R^{S_N}$. $\mathcal{Q}$ is therefore the centered product Gaussian probability on Ω of variance N. By duality, S_N acts on Ω through the rule $(\sigma.h)(\sigma') \equiv h(\sigma.\sigma')$, where $\sigma, \sigma' \in S_N$ and $h \in \Omega$.
For each choice of $H \in \Omega$, let us denote by X^H the Metropolis dynamics with Hamiltonian H, i.e. X^H is the Markov process with generator

$$L^H f(\sigma) \equiv \frac{1}{N} \sum_{i=1}^{N} \epsilon^{-\beta[H(\underline{i}.\sigma)-H(\sigma)]^+} \left(f(\underline{i}.\sigma) - f(\sigma)\right) \tag{5.1}$$

here, as before $[x]^+$ is the positive part of x. We denote by $P_t^H \equiv e^{tL^H}$, its semi-group, and let $I\!E_\sigma^H$ be the law of X^H when $X^H(0) = \sigma$.
Let us now define the stochastic process $\omega_t \equiv X_t^H.H$. The state space of ω_t is Ω. ω_t is simply the Hamiltonian translated according to the position of the particle. For instance note that, by definition, $\omega_t(1\!\!1) = X_t^H.H(1\!\!1) = H(X_t^H)$ is nothing but the value of the Hamiltonian evaluated at the position of the particle at time t. We consider the canonical construction of the Markov process, X_t^H, so we call X_t the coordinate process on the space of cad-lag functions taking value in S_N. We call $I\!E_\sigma^H$, the law of the Markov process with generator L^H starting from σ. We denote e_σ^H, the law of the process $\omega_t \equiv X_t.H$ when X_t is distributed according to $I\!E_\sigma^H$.
By definition we have

$$e_\sigma^H[\phi_1(\omega_{t_1})...\phi_k(\omega_{t_k})] = I\!E_\sigma^H[\phi_1(X_{t_1}.H)...\phi_k(X_{t_k}.H)] \tag{5.2}$$

The point is the following

Lemma 5.1

$$e_\sigma^H = e_1^{\sigma.H} \tag{5.3}$$

Proof:
Let ϕ be some measurable function on Ω. Define $\phi^H(\sigma) \equiv \phi(\sigma.H)$. Note that

$$L^{\sigma.H}\phi^{\sigma.H}(\sigma') = L^H\phi^H(\sigma.\sigma')$$

this follows from

$$\begin{aligned}
\left(L_N^{\sigma.H} f^{\sigma.H}\right)(\sigma') &= \sum_{i=1}^{N} e^{-\beta[\sigma.H(\underline{i}.\sigma')-\sigma.H(\sigma')]^+} \Big(\phi\big((\underline{i}.\sigma').\sigma.H\big) - \phi\big(\sigma'.\sigma.H\big)\Big) \\
&= \sum_{i=1}^{N} e^{-\beta[H(\underline{i}.(\sigma.\sigma')-H(\sigma.\sigma')]^+} \Big(\phi\big(\underline{i}.(\sigma.\sigma').H\big) - \phi\big((\sigma.\sigma').H\big)\Big) \\
&= \left(L_N^H\phi^H\right)(\sigma.\sigma')
\end{aligned} \tag{5.4}$$

Therefore, since $\phi^{\sigma.H}(\sigma') = \phi^H(\sigma.\sigma')$, we get

$$\left(e^{tL_N^{\sigma.H}}\phi^{\sigma.H}\right)(\sigma') = \left(e^{tL_N^H}\phi^H\right)(\sigma.\sigma') \tag{5.5}$$

Applying this last equality for $\sigma' = \mathbb{1}$, we have proved that

$$e_{\mathbb{1}}^{\sigma.H}\left[\phi(\omega_t)\right] = e_\sigma^H\left[\phi(\omega_t)\right] \tag{5.6}$$

that is (5.3) holds for functions of one coordinate.
To extend it to an arbitrary cylindrical function we have, assuming $t_1 < t_2$

$$\begin{aligned} e_\sigma^H\left[\phi_1(\omega_{t_1})\phi_2(\omega_{t_2})\right] &= \mathbb{E}_\sigma^H\left[\phi_1(X_{t_1}.H)\phi_2(X_{t_2}.H)\right] \\ &= \mathbb{E}_\sigma^H\left[\phi_1(X_{t_1}.H)\mathbb{E}_{X_{t_1}}\left[\phi_2(X_{t_2-t_1}.H)\right]\right] \end{aligned} \tag{5.7}$$

where at the last step we have used that $X_t \equiv X_t^H$ is an homogeneous Markov process. Using (5.3), we have

$$\mathbb{E}_{X_{t_1}}^H\left[\phi_2(X_{t_2-t_1}.H)\right] = \mathbb{E}_{\mathbb{1}}^{X_{t_1}.H}\left[\phi_2(X_{t_2-t_1}.X_{t_1}.H)\right] \tag{5.8}$$

using again (5.3) twice, we have also

$$\begin{aligned} &\mathbb{E}_\sigma^H\left[\phi_1(X_{t_1}.H)\mathbb{E}_{\mathbb{1}}^{X_{t_1}.H}\left[\phi_2(X_{t_2-t_1}.X_{t_1}.H)\right]\right] \\ &= \mathbb{E}_{\mathbb{1}}^{\sigma.H}\left[\phi_1(X_{t_1}.\sigma.H)\mathbb{E}_{\mathbb{1}}^{X_{t_1}.\sigma.H}\left[\phi_2(X_{t_2-t_1}.X_{t_1}.\sigma.H)\right]\right] \\ &= \mathbb{E}_{\mathbb{1}}^{\sigma.H}\left[\phi_1(X_{t_1}.\sigma.H)\mathbb{E}_{X_{t_1}}^{\sigma.H}\left[\phi_2(X_{t_2-t_1}.\sigma.H)\right]\right] \end{aligned} \tag{5.9}$$

Using once again the Markov property for X_t, we get

$$\begin{aligned} &\mathbb{E}_{\mathbb{1}}^{\sigma.H}\left[\phi_1(X_{t_1}.\sigma.H)\mathbb{E}_{X_{t_1}}^{\sigma.H}\left[\phi_2(X_{t_2-t_1}.\sigma.H)\right]\right] \\ &= \mathbb{E}_{\mathbb{1}}^{\sigma.H}\left[\phi_1(X_{t_1}.\sigma.H)\phi_2(X_{t_2}.\sigma.H)\right] \\ &= e_{\mathbb{1}}^{\sigma.H}\left[\phi_1(\omega_{t_1})\phi_2(\omega_{t_2})\right] \end{aligned} \tag{5.10}$$

Now it is easy to generalize what we just did to an arbitrary product of functions of one coordinate, then to cylindrical function and to measurable function by the monotone class theorem. This ends the proof of the lemma. ∎

Note that ω_t is the image of X_t by the map $X_t \to X_t.H$. In general the image of a Markov process is not Markovian, however here we have the

Lemma 5.2 *ω_t is an homogeneous Markov process.*

Proof: It is enough to prove that

$$e_1^{\sigma.H}\left[\phi_1(\omega_{t_1})\phi_2(\omega_{t_2})\right] = e_1^{\sigma.H}\left[\phi_1(\omega_{t_1})e_1^{\omega_{t_1}}\left[\phi_2(\omega_{t_2-t_1})\right]\right] \tag{5.11}$$

We have

$$e_1^{\sigma.H}\left[\phi_1(\omega_{t_1})\phi_2(\omega_{t_2})\right] = I\!E_\sigma^H\left[\phi_1(X_{t_1}.H)\phi_2(X_{t_2}.H)\right] \tag{5.12}$$

Since X_t is an homogeneous Markov process, we have

$$I\!P_\sigma^H\left[X_{t_1}=\sigma_1, X_{t_2}=\sigma_2\right] = I\!P_\sigma^H\left[X_{t_1}=\sigma_1\right] I\!P_{\sigma_1}^H\left[X_{t_2-t_1}=\sigma_2\right] \tag{5.13}$$

Therefore we get

$$\begin{aligned} &I\!E_\sigma^H\left[\phi_1(X_{t_1}.H)\phi_2(X_{t_2}.H)\right] = \\ &\quad\sum_{H_1'}\phi_1(H_1')\sum_{\sigma_1}1\!\!I_{\{\sigma_1.H=H_1'\}} I\!P_\sigma^H\left[X_{t_1}=\sigma_1\right] I\!E_{\sigma_1}^H\left[\phi_2(X_{t_2-t_1}.H)\right] \end{aligned} \tag{5.14}$$

The point is that using (5.3), we have

$$\begin{aligned} &\sum_{\sigma_1}1\!\!I_{\{\sigma_1.H=H_1'\}} I\!P_\sigma^H\left[X_{t_1}=\sigma_1\right] I\!E_{\sigma_1}^H\left[\phi_2(X_{t_2-t_1}.H)\right] = \\ &\qquad e_1^{H_1'}\left[\phi_2(\omega_{t_2-t_1})\right]\sum_{\sigma_1}1\!\!I_{\{\sigma_1.H=H_1'\}} I\!P_\sigma^H\left[X_{t_1}=\sigma_1\right] \end{aligned} \tag{5.15}$$

Therefore we get

$$\begin{aligned} &I\!E_\sigma^H\left[\phi_1(X_{t_1}.H)\phi_2(X_{t_2}.H)\right] \\ &\quad= \sum_{H_1'}\phi_1(H_1')e_1^{H_1'}\left[\phi_2(\omega_{t_2-t_1})\right]\sum_{\sigma_1}1\!\!I_{\{\sigma_1.H=H_1'\}} I\!P_\sigma^H\left[X_{t_1}=\sigma_1\right] \\ &\quad= I\!E_\sigma^H\left[\phi_1(X_{t_1}.H)e_1^{X_{t_1}.H}\left[\phi_2(\omega_{t_2-t_1})\right]\right] \\ &\quad= e_1^{\sigma.H}\left[\phi_1(\omega_{t_1})e_1^{\omega_{t_1}}\left[\phi_2(\omega_{t_2-t_1})\right]\right] \end{aligned} \tag{5.16}$$

which is what we wanted to prove. ■

Let π_β^H be the Gibbs measure with Hamiltonian H i.e.

$$\pi_\beta^H(\sigma) \equiv \frac{e^{-\beta H(\sigma)}}{Z^H(\beta)}$$

and let us define the probability ν_β^H on Ω by

$$\nu_\beta^H(f) \equiv \sum_{\sigma\in S_N} f(\sigma.H)\pi_\beta^H(\sigma) \tag{5.17}$$

when $f:\Omega\to I\!R$. That is for all $H'\in\Omega$

$$\nu_\beta^H(H') = \sum_{\sigma\in S_N} \pi_\beta^H(\sigma)1\!\!1_{\{H'=\sigma.H\}} \tag{5.18}$$

We have the

Lemma 5.3 *For each $H\in\Omega$, ν_β^H is an invariant and reversible measure for ω_t*

Proof:
The invariance follows from

$$\begin{aligned}\sum_{H'}\nu_\beta^H(H')e_1^{H'}\left[\phi(\omega_t)\right] &= \sum_{H'}\sum_{\sigma\in\Sigma_N}\pi_\beta^H(\sigma)1\!\!1_{\{H'=\sigma.H\}}e_1^{H'}\left[\phi(\omega_t)\right]\\ &= \sum_{\sigma\in S_N}\pi_\beta^H(\sigma)I\!\!E_\sigma^H\left[\phi^H(X_t)\right]\\ &= \sum_{\sigma\in S_N}\pi_\beta^H(\sigma)\phi^H(\sigma)\\ &= \nu_\beta^H(\phi)\end{aligned} \tag{5.19}$$

where we have used the fact that π_β^H is invariant for X_t at the third step.
The reversibility follows from

$$\sum_{\sigma\in S_N}\phi^H(\sigma)\pi_\beta^H(\sigma)I\!\!E_\sigma^H\left[\psi^H(X_t)\right] = \sum_{\sigma\in S_N}\psi^H(\sigma)\pi_\beta^H(\sigma)I\!\!E_\sigma^H\left[\phi^H(X_t)\right] \tag{5.20}$$

since π_β^H is reversible for X_t. This ends the proof of the lemma. ■

Now, for any bounded measurable function f defined on Ω, we have, as t tends to $+\infty$,

$$e_\sigma^H[f(\omega_t)] = I\!\!E_\sigma^H[f(X_t^H.H)] \to \nu_\beta^H(f) \tag{5.21}$$

We are interested in estimating the speed on convergence in (5.21). A fundamental fact is stated in the following lemma

Lemma 5.4 *For any $\varphi:\Omega_N\to I\!R$, $\mathcal{Q}\left[e_\sigma^H(\varphi(\omega_t))\right]$ is independent of $\sigma\in S_N$*

Proof: This follows from the fact that on the one hand, for all $\sigma\in S_N$ and for all $f:\Omega\to I\!R$, we have

$$\mathcal{Q}\left[f(H)\right] = \mathcal{Q}\left[f(\sigma H)\right] \tag{5.22}$$

since $\mathcal{Q}$ is invariant by any permutation of the configurations H.
Therefore, using (5.4), we have, for all $\varphi : \Omega \to I\!R$

$$\mathcal{Q}\left[e_{\sigma}^{H}(\phi)\right] = \mathcal{Q}\left[e_{1}^{\sigma.H}(\phi)\right] = \mathcal{Q}\left[e_{1}^{H}(\phi)\right] \tag{5.23}$$

which is what we wanted to prove. ■

Now we can define the following time:

$$T_{av}(\epsilon) \equiv \inf\left\{t > 0 \;\; s.t. \; \sup_{s\geq t} \sup_{\varphi:\|\varphi\|_\infty \leq 1} \mathcal{Q}\left[\left|e_{\sigma}^{H}(\varphi(\omega_t)) - \nu_{\beta}^{H}(\varphi)\right|\right] \leq \epsilon\right\} \tag{5.24}$$

here $\|\varphi\|_\infty = \sup_{\omega\in\Omega} |\varphi(\omega)|$. $T_{av}(\epsilon)$ is the time such that the average over the medium of the medium as seen from the process is definitively within ϵ of the reversible measure ν_{β}^{H}.
The main result of this section is

Theorem 5.5 *For all $\epsilon > 0$, for all $\beta \leq \beta_c$,*

$$\limsup_{N\uparrow\infty} \frac{1}{N} \log T_{av}(\epsilon) \leq \beta^2 \tag{5.25}$$

Proof: Using Lemma 5.4, denoting by $d\eta(x)$ the uniform measure on S_N, we get

$$\mathcal{Q}\left[\left|e_{\sigma}^{H}(\varphi(\omega_t)) - \nu_{\beta}^{H}(\varphi)\right|\right] = \int d\eta(\sigma)\mathcal{Q}\left[\left|e_{\sigma}^{H}(\varphi(\omega_t)) - \nu_{\beta}^{H}(\varphi)\right|\right] \tag{5.26}$$

since the left hand side does not depends on σ. Now using Tonelli's theorem we get

$$\int d\eta(\sigma)\mathcal{Q}\left[\left|e_{\sigma}^{H}(\varphi(\omega_t)) - \nu_{\beta}^{H}(\varphi)\right|\right] = \mathcal{Q}\left[\int d\eta(\sigma)\left|e_{\sigma}^{H}(\varphi(\omega_t)) - \nu_{\beta}^{H}(\varphi)\right|\right] \tag{5.27}$$

Now, since

$$e_{\sigma}^{H}(\varphi(\omega_t)) = I\!E_{\sigma}^{H}(\varphi^{H}(X_t)) \tag{5.28}$$

using (5.17), we get

$$\nu_{\beta}^{H}(\varphi) = \pi_{\beta}^{H}(\varphi^{H}) \tag{5.29}$$

Therefore

$$\begin{aligned} e_{\sigma}^{H}(\varphi(\omega_t)) - \nu_{\beta}^{H}(\varphi) &= I\!E_{\sigma}^{H}(\varphi^{H}(X_t)) - \pi_{\beta}^{H}(\varphi^{H}) \\ &= \left(P_{t}^{H}(\varphi^{H})\right)(\sigma) - \pi_{\beta}^{H}(\varphi^{H}) \end{aligned} \tag{5.30}$$

therefore collecting what we just did we get

$$\mathcal{Q}\left[\left|e_{\sigma}^{H}(\varphi(\omega_t)) - \nu_{\beta}^{H}(\varphi)\right|\right] = \mathcal{Q}\left[\int d\eta(\sigma')\left|\left(P_{t}^{H}(\varphi^{H})\right)(\sigma') - \pi_{\beta}^{H}(\varphi^{H})\right|\right] \tag{5.31}$$

To continue, recalling Proposition 4.7 and 4.9, for all c, let $\mathcal{A}(c)$ be the subspace of Ω, of $\mathcal{Q}$-probability bigger than $1-2e^{-cN}$, which is the intersection of the two subspaces where we have the estimates (4.32) and (4.58).
Then we get

$$\begin{aligned}\mathcal{Q}\left[\int d\eta(x)\left|\left(P_t^H(\varphi^H)\right)(x)-\pi_\beta^H(\varphi^H)\right|\right] &\leq 2\mathcal{Q}\left[\|\varphi\|_\infty \mathbb{1}_{\mathcal{A}^c(c)}\right]+\\ &+C_p t^{-p/(2-p)}\mathcal{Q}\left[\mathbb{1}_{\mathcal{A}(c)}\|\varphi^H\|_\infty(\mathcal{L}_\eta^H(p))^{-p/2}(\mathcal{L}^H(p))^{-p^2/(4-2p)}\right]\end{aligned} \tag{5.32}$$

where the first part of the inequality follows from the fact that for all $H\in\Omega$ and all $t>0$, P_t^H is a contraction operator from $L^\infty[\Omega, I\!R]$ into itself and π_β^H is a probability measure. The second part follows from (2.11). We recall that $C_p=e^{-p/2}2^{p/2}((2-p)/p)^{-p^2/(4-2p)}$. Using now Proposition 4.7, Proposition 4.9 and $\|\varphi\|_\infty\leq 1$, we get

$$\begin{aligned}&\mathcal{Q}\left[\left|e_\sigma^H(\varphi(\omega_t))-\nu_\beta^H(\varphi)\right|\right]\leq 2e^{-cN}+\\ &+C_p t^{-p/(2-p)}(22)^{p^2/(4-2p)}(4)^{p/2}(N)^{8p/4-2p}(e^{2\beta\beta_c\sqrt{c_1(1+c)N\log N}}e^{\beta dN})^{p((4-p)/4-2p)}\end{aligned} \tag{5.33}$$

From now on the proof is exactly the same as the proof of Theorem 4.1.
At this point it is clear that we can also gives estimates that are similar to the ones given in Theorem 4.2 by using the same arguments as before and the computation done in the proof of Theorem 4.2. Let us state it as a Theorem.

Theorem 5.6 *For all N large enough, for all $\beta\leq\beta_c$, There exists a constant $c_1>0$, such that for all $c>0$, there exists a constant $C_0=C_0(c,\beta)$ such that*

$$\begin{aligned}&\frac{1}{N}\log T_{av}(e^{-N^{1/4}(\log N)^{3/4}})\\ &\leq\beta^2+2\beta\beta_c\left(\frac{c_1(1+c)\log N}{N}\right)^{1/2}+c_2(\beta,c)\left(\frac{\log N}{N}\right)^{1/4}+C_0(\frac{\log N}{N})^{3/4}\end{aligned} \tag{5.34}$$

where

$$c_2(\beta,c)\equiv\beta\left(12\beta\beta_c\sqrt{c_1(1+c)}\right)^{1/2}+\frac{1}{4}\frac{\beta^2+\beta_c^2}{\beta\beta_c\sqrt{c_1(1+c)}} \tag{5.35}$$

Moreover for all $\delta>0$

$$\begin{aligned}&\frac{1}{N}\log T_{av}(N^{-\delta})\leq\beta^2+\beta\left(12\beta\beta_c\sqrt{c_1(1+c)}\right)^{1/2}(\frac{\log N}{N})^{1/4}\\ &+2\beta\beta_c\left(\frac{c_1(1+c)\log N}{N}\right)^{1/2}+\frac{1}{4}\frac{\beta^2+\beta_c^2}{\beta\beta_c}\frac{\delta}{\sqrt{c_1(1+c)}}(\frac{\log N}{N})^{1/2}+C_0\frac{\log N}{N}\end{aligned} \tag{5.36}$$

VI. Statics estimates for the REM

In this section we will give some estimates for the various constrained partition functions and partition functions on small spaces for the REM. These are just adaptations of similar estimates done in [4] section 4.2.1.

Let us first prove Lemma 4.8. We denote by $Z_\alpha(\beta, \geq -d) \equiv Z_{j-1}(\beta, \geq -d)[z_j, z_{>j}]$. Let M be as in Lemma 4.8, and make the partition of the real interval $(-\infty, dN]$ with the intervals

$$\Delta_0 \equiv \left(-\infty, \beta_c \frac{N}{M}\right] \tag{6.1}$$

if $1 \leq k \leq \frac{d}{\beta_c} M - 1$

$$\Delta_k \equiv \left(\beta_c \frac{k}{M} N, \beta_c \frac{k+1}{M}\right] \tag{6.2}$$

Let

$$N_k = N_k(z_j, z_{>j}) = \sum_{x \in \{-1,+1\}^{j-1}} \mathbb{I}_{\Delta_k}(-H(x, z_j, z_{>j})) \tag{6.3}$$

be the occupation number of the interval Δ_k, it is easy to check that, if $p_k = \mathbb{P}\left[-H(x) \in \Delta_k\right]$, then

$$\beta_c \frac{\sqrt{N}}{M} 2^{-\frac{(k+1)^2}{M^2} N} < p_k < \beta_c \frac{\sqrt{N}}{M} 2^{-\frac{k^2}{M^2} N} \tag{6.4}$$

Using the exponential Markov inequality and optimizing we get

$$\mathbb{P}\left[N_k > \rho_k \mathbb{E}(N_k)\right] \leq \exp\left\{-\lambda_k 2^{\alpha N}\right\} \tag{6.5}$$

where

$$\rho_k = 2^{N[\frac{(k+1)^2}{M^2} - \alpha]^+ + 2} \tag{6.6}$$

and if $\rho_k p_k \geq 1$, $\lambda_k = \infty$, while if $\rho_k p_k < 1$

$$\lambda_k \equiv \rho_k p_k \log \frac{\rho_k (1 - p_k)}{1 - \rho_k p_k} - \log\left[1 - p_k + \frac{\rho_k p_k (1 - p_k)}{1 - \rho_k p_k}\right] \tag{6.7}$$

It is not too long to check that $\lambda_k \geq \rho_k p_k c_1$ for some positive constant c_1, and also $\rho_k p_k \geq 2^{N/M^2}$, therefore with our choice of M, we get

$$\begin{aligned} \mathbb{P}\left[N_k > \rho_k \mathbb{E}(N_k)\right] &\leq \exp -\left\{c_1 2^{N/M^2}\right\} \\ &\leq 2^{-2N} \exp -(cN) \end{aligned} \tag{6.8}$$

Note that the term 2^{-2N} will be more than enough to get uniformity with respect to the index i for the chosen family of path, the index j, the configurations $z_j, z_{>j}$, and the index k.
Therefore, calling $A \equiv \sqrt{\alpha M^2 - 1}$ and $D+1 \equiv dM/\beta_c$ and using (6.4) and (6.6), we get

$$\begin{aligned} Z_\alpha(\beta, \geq -d) \leq 2^j e^{\beta\beta_c \frac{N}{M}} + \sum_{k=1}^{A\wedge D} \frac{\sqrt{N}}{M} e^{N(\alpha - \frac{k^2}{M^2})\frac{\beta_c^2}{2} + \beta\beta_c \frac{(k+1)}{M} N} \\ + \sum_{k=A\wedge D+1}^{D} \frac{\sqrt{N}}{M} e^{N\beta\beta_c \frac{(K+1)M}{N}} 2^{N/M^2} \end{aligned} \tag{6.9}$$

where the last sum is not present if $D < A$.
We have

$$2^j e^{\beta\beta_c \frac{N}{M}} \leq 2^J e^{\beta\beta_c \sqrt{c_1(1+c)N\log N}} \tag{6.10}$$

It is immediate to see that, if $D < A$

$$\sum_{k=A\wedge D+1}^{D} \frac{\sqrt{N}}{M} e^{N\beta\beta_c \frac{(K+1)M}{N}} 2^{N/M^2} \leq c_u N^{3/2} e^{\beta d N} \tag{6.11}$$

It remains to estimate the first sum in the right hand side of (6.9). Let us call it $S(N)$, if we denote $x = K/M$, the maximum in the exponent occurs for $x = \beta/\beta_c$. Therefore, if $A < \frac{\beta}{\beta_c} M$ we easily get

$$\begin{aligned} S(N) &\leq \sqrt{N} e^{\beta\beta_c \sqrt{c_1(1+c)N\log N}} e^{N\beta\beta_c\sqrt{\alpha}} \\ &\leq \sqrt{N} e^{\beta\beta_c \sqrt{c_1(1+c)N\log N}} e^{N\beta d} \end{aligned} \tag{6.12}$$

where at the last step we have used $\beta_c\sqrt{\alpha} < \beta < d$
If $\frac{\beta}{\beta_c} M \leq A < D$ we easily get

$$S(N) \leq \sqrt{N} e^{\beta\beta_c \sqrt{c_1(1+c)N\log N}} e^{N(\frac{\beta^2}{2} + \alpha\frac{\beta_c^2}{2})} \tag{6.13}$$

If $D \leq A$, since $d > \beta$, the maximum of the exponent occurs inside the interval of summation therefore we easily get that

$$S(N) \leq \sqrt{N} e^{\beta\beta_c \sqrt{c_1(1+c)N\log N}} e^{N(\frac{\beta^2}{2} + \alpha\frac{\beta_c^2}{2})} \tag{6.14}$$

collecting (6.10) to (6.14) we get (4.52) and (4.53).
The Lemma 4.6 is proved in exactly the same way, by making a similar partition of $[dN, +\infty)$, for proving (4.26). Restricting the sum over k to just the one corresponding to $k = M\beta/\beta_c$, it is easy to get (4.27).

References

[1] Derrida B. (1980) Random Energy model: Limit of a family of disordered systems. *Phys. Rev. Lett* **45** 79–82.

[2] Derrida B. (1981) Random Energy model: An exactly solvable model of a spin glass. *Phys. Rev. B* **24** 2613–2626.

[3] Eisele, T. (1983) On a Third Order Phase Transition. *Comm. Math. Phys.* **90**, 125–159.

[4] Fontes, L.R.G., Isopi, M, Kohayakawa, Y., and Picco, P. (1998) The sprectral gap of the REM under metroplolis dynamics. *Ann. Appl. Prob* **8** 917–943.

[5] Mathieu P. (1997) Hitting times and spectral gap inequalities. *Ann. Instit. H.Poincaré. Série Probabilités* **33**. n 4, 437–465.

[6] Mathieu P. (1999) Sur la convergence des marches aléatoires dans un milieu aléatoire et les inégalités de Poincaré généralisées. *CRAS Paris*, **t. 329, série I**, 1015–1020

[7] Mathieu P. (2000) Convergence to equilibrium for spin glasses. *Comm. Math. Phys.* **215**, 57–68.

[8] Olivieri E., Picco P. (1984) On the Existence of Thermodynamics for the Random Energy Model. *Comm. Math. Phys.* **96**, 125–144.

[9] Galves.A., Martinez S., and Picco P. (1989) Fluctuations in Derrida's Random Energy and Generalized Random Energy Models. *J. Stat. Phys.* **54** 515–529.

[10] Mezard M., Parisi G., and Virasoro M.,(1980) *Spin Glass theory and beyond*, World Scientific.

[11] Saloff-Coste L. (1997). *Lectures on finite Markov chains.* Ecole d'été de St-Flour. Lectures Notes in Math. Springer Verlag, Berlin.

Pierre Mathieu
CMI, Université de Provence,
39 Rue F. Joliot Curie, 13453 Marseille Cedex 13, France.
pmathieu@gyptis.univ-mrs.fr

Pierre Picco
CPT.CNRS Luminy, case 907,
13288 Marseille Cedex 9, France.
Picco@cpt.univ-mrs.fr.